ALGEBRA LINEAL

Juan de Burgos

Mc Graw Hill

ALGEBRA LINEAL

ALGEBRA LINEAL

Juan de Burgos Román

Catedrático de la Escuela Superior de
Ingenieros Aeronáuticos
Universidad Politécnica de Madrid

Esta obra ha sido galardonada con el premio al mejor libro de texto para escuela o facultad de la Universidad Politécnica de Madrid. 1994.

McGraw-Hill

MADRID · BUENOS AIRES · CARACAS · GUATEMALA · LISBOA · MEXICO
NUEVA YORK · PANAMA · SAN JUAN · SANTAFE DE BOGOTA · SANTIAGO · SAO PAULO
AUCKLAND · HAMBURGO · LONDRES · MILAN · MONTREAL · NUEVA DELHI · PARIS
SAN FRANCISCO · SIDNEY · SINGAPUR · ST. LOUIS · TOKIO · TORONTO

ALGEBRA LINEAL

No está permitida la reproducción total o parcial de este libro, ni su tratamiento informático, ni la transmisión de ninguna forma o por cualquier medio, ya sea electrónico, mecánico, por fotocopia, por registro u otros métodos, sin el permiso previo y por escrito de los titulares del Copyright.

DERECHOS RESERVADOS © 1993, respecto a la primera edición en español por

McGRAW-HILL/INTERAMERICANA DE ESPAÑA, S. A.

Edificio Valrealty A, 1.ª planta
Basauri, 17
28023 ARAVACA (Madrid)

ISBN: 84-481-0134-0
Depósito legal: M. 12.580-1995

Editor: **Juan** Stumpf
Cubierta: Diseño de Félix Piñuela. Grafismo electrónico, basado en el manuscrito de Abdalla Benerbocar del siglo XIII perteneciente a la Biblioteca del Monasterio de El Escorial
Compuesto en: Fernández Ciudad, S. L.
Impreso por: Impresos y Revistas, S. A. (IMPRESA)

IMPRESO EN ESPAÑA - PRINTED IN SPAIN

A mi madre, Rosario Román,
en reconocimiento a su admirable
quehacer docente.

Contenido

PROLOGO xiii

PARTE I

INICIACION: LINEALIDAD Y RANGO

Introducción 2

Capítulo 1. Sistemas de ecuaciones lineales; el método de Gauss 3

α *Primeras definiciones. Equivalencia* 3
1.1. Sistemas de ecuaciones lineales 3
1.2. Sistemas equivalentes 8

β *El método de Gauss* 12
1.3. Matrices escalonadas y sistemas escalonados 12
1.4. El método de Gauss 16
1.5. Sistemas homogéneos con menos ecuaciones que incógnitas 21

Capítulo 2. Rango (de vectores y de matrices) 22

α *Vectores de n componentes* 22
2.1. El espacio vectorial K^n 23
2.2. Dependencia e independencia lineal 26

β *Rango de un sistema de vectores* 30
2.3. Rango. Operaciones elementales 30
2.4. Cálculo del rango de un sistema de vectores 34

γ *Rango de una matriz* 36
2.5. Existencia y cálculo del rango 36
2.6. Matrices equivalentes 40

Capítulo 3. Operaciones con matrices; matriz inversa 42

α *Matrices; álgebra de matrices* 43
3.1. Primeras definiciones 43
3.2. Producto de matrices 46
3.3. Traspuesta de una matriz 54
3.4. Relación entre las operaciones elementales y el producto 55
3.5. Multiplicación de matrices por bloques 60

β *Matrices invertibles* 62
3.6. Definición y primeras propiedades 62
3.7. Cálculo efectivo de la inversa 65
3.8. Caracterizaciones de las matrices invertibles 67

Capítulo 4. Determinantes **70**
α Definición y valor de un determinante **71**
4.1. Determinante de una matriz cuadrada 72
4.2. Expresión del valor de un determinante 75
4.3. Cálculo efectivo de un determinante 83
β Propiedades: desarrollo e inversa **85**
4.4. Determinante de un producto: consecuencia 85
4.5. Desarrollo por los elementos de una línea 87
4.6. Matriz inversa 91
4.7. Rango de menores de una matriz 93
γ Teoremas de Cramer y de Rouché **96**
4.8. Sistemas de Cramer 98
4.9. Teorema de Rouché 99
Ejercicios y problemas a la parte I **103**
Enunciados 103
Soluciones 111

PARTE II

ALGEBRA LINEAL

Capítulo 5. Espacios vectoriales **116**
α Espacios, subespacios y combinaciones lineales **116**
5.1. Concepto de espacio vectorial 116
5.2. Subespacios vectoriales 118
5.3. Dependencia e independencia lineal 120
β Bases. Coordenadas **126**
5.4. Espacios de dimensión finita 126
5.5. Coordenadas 132
5.6. Rango de un sistema de vectores 137
γ Suma de subespacios **139**
5.7. Suma y suma directa 139
5.8. Dimensión de la suma 146

Capítulo 6. Aplicaciones lineales **149**
α Aplicaciones lineales **149**
6.1. Definición y propiedades 149
6.2. Isomorfismos 156
β Matrices de las aplicaciones lineales **160**
6.3. Ecuaciones y matriz de una aplicación lineal 160
6.4. Matrices equivalentes 167
γ Operaciones con aplicaciones lineales **175**
6.5. Espacios vectoriales de homomorfismos y de matrices 176
6.6. Anillo de endomorfismos. Matrices invertibles 179
δ Sistemas de ecuaciones lineales **185**
6.7. Variedades afines 186
6.8. Teorema de Rouché 188
ε Espacio dual **190**
6.9. Espacio dual 190
6.10. Aplicación dual o traspuesta 193

Ejercicios y problemas a la parte II **196**
Enunciados 196
Soluciones 202

PARTE III

Capítulo 7. Formas cuadráticas **208**

α Formas cuadráticas. Conjugación **208**
7.1. Formas bilineales 208
7.2. Formas cuadráticas 218
7.3. Conjugación respecto de una forma cuadrática 224

β Diagonalización de una forma cuadrática **231**
7.4. Forma diagonal (congruencia) 231
7.5. Diagonalización efectiva de una forma cuadrática 234

γ Formas cuadráticas reales **239**
7.6. Formas definidas y ley de inercia 240
7.7. Expresión canónica de una forma cuadrática 246

Capítulo 8. Espacios vectoriales euclídeos **253**

α Producto escalar **254**
8.1. Producto escalar de vectores 254
8.2. Normas y ángulos 259

β Vectores ortogonales y ortonormales **266**
8.3. Bases ortonormales 267
8.4. Proyección ortogonal 277

γ Transformaciones ortogonales **284**
8.5. Transformaciones ortogonales 285
8.6. Matrices ortogonales 290
8.7. Transformaciones ortogonales en 2 y 3 dimensiones 296

δ Producto mixto y producto vectorial **303**
8.8. Producto mixto 304
8.9. Producto vectorial 307

Ejercicios y problemas a la parte III **313**
Enunciados 313
Soluciones 320

PARTE IV

Capítulo 9. Diagonalización de endomorfismos y de matrices **326**

α Autovalores y autovectores **326**
9.1. Autovalores y autovectores. Polinomio característico 326
9.2. Multiplicidades algebraica y geométrica de un autovalor 336
9.3. Complejificación de un espacio vectorial real 341

β Endomorfismos diagonalizables **345**
9.4. Diagonalización (por semejanza) 345
9.5. Diagonalización ortogonal 353

γ Diagonalización por bloques de una transformación ortogonal **363**
9.6. Subespacios invariantes de una transformación ortogonal 364
9.7. Forma canónica de una transformación ortogonal 369

Capítulo 10. Forma canónica de Jordan **376**

α Forma de Jordan de una matriz **377**
10.1. Diagonalización en bloques triangulares 377
10.2. Forma canónica de Jordan 384

β Obtención de la forma de Jordan **398**
10.3. Tamaño de los bloques de Jordan 399
10.4. Matriz de paso a la forma de Jordan 402

Ejercicios y problemas a la parte IV **412**
Enunciados 412
Soluciones 418

PARTE V

GEOMETRIA CARTESIANA

Capítulo 11. Los espacios geométricos E_2 y E_3 ..... **426**

α Axiomas y definiciones **426**
11.1. Los espacios geométricos bi y tridimensionales 427
11.2. Las rectas y los planos 430

β Geometría plana (afín y euclídea) **433**
11.3. Problemas afines (en el plano E_2) 433
11.4. Problemas euclídeos (en el plano E_2) 441

γ Geometría tridimensional (afín y euclídea) **452**
11.5. Ecuaciones de rectas y planos 452
11.6. Posiciones relativas de rectas y planos 459
11.7. Angulos y distancias en E_3 470

δ Espacio afín ampliado E_n ($n=2$ y $n=3$) **485**
11.8. Coordenadas homogéneas 485
11.9. Espacio afín ampliado (o proyectivizado) 491

Capítulo 12. Espacios puntuales (caso general) **498**

α Espacios afín y afín euclídeo **498**
12.1. Los espacios puntuales 499
12.2. Coordenadas cartesianas 505
12.3. Variedades lineales afines 509
12.4. Variedades lineales de dimensión finita 512

β Cuestiones de carácter afín **518**
12.5. Intersección y suma de variedades lineales 518
12.6. Coordenadas baricéntricas 522

γ Cuestiones de carácter euclídeo **527**
12.7. Ortogonalidad y mínima distancia 527
12.8. La esfera 533

Ejercicios y problemas a la parte V **540**
Enunciados 540
Soluciones 546

PARTE VI

Capítulo 13. Cónicas y cuádricas: estudio particular **552**

α Estudio particular de cónicas **552**

Prólogo

Es aquí, en el Prólogo, donde los autores, después de contemplar la obra que acaban de escribir, se ponen a hablar de ella a los demás, la «presentan en sociedad». Y lo harán, aun sin buscarlo, por muy ecuánimes que quieran ser, señalando sus virtudes, que con alguna contará, achicando sus defectos, que tampoco han de faltar, justificando sus carencias, realzando su importancia, tratándola con el natural apego que cualquiera siente por todo aquello que hace, fabrica, produce, inventa o imagina. Al hilo de ello, bueno es reparar, por otra parte, en que, de no hacerse las cosas como se ha dicho, el lector se dirá: «si a éste, que la ha escrito, no le satisface, ¿a quién le ha de gustar la obra?».

Es obligado decir ahora, en primer lugar, que este libro, que ha terminado llamándose «Algebra lineal», ha estado a punto de titularse «Algebra lineal y geometría cartesiana». Y es que un tercio, al menos, de su contenido versa sobre geometría. Entendiendo, no obstante, que la geometría, la que aquí se estudia, se sustenta en el álgebra, de la que hereda y utiliza conceptos, métodos y modos de hacer, se ha creído que, diciendo sólo «álgebra lineal», también se decía, implícitamente, «geometría cartesiana».

En esto de presentar el libro a sus lectores, un asunto obligado es el de hacer la pertinente glosa acerca de los contenidos. Sobre este particular, iremos a salto de mata, como en todo lo demás; hablaremos sólo de aquello que pueda ser chocante o inusual y que, por tanto, convenga explicar, de aquello que precise de aclaración o comentario. Al que le urja conocer con detalle las materias que se abordan en el texto, habrá acudido ya al índice de contenidos, donde todo figura en su debido orden. No está de más, aunque tampoco estaría de menos, decir aquí que, en lo tocante a ser leídos, me malicio que los contenidos salen mejor parados que los prólogos: según tengo observado, lo primero que, normalmente, hace quien hojea un texto es dirigirse a la tabla de Contenido; al Prólogo, por el contrario, sólo van los lectores, si es que lo hacen, después de haber leído una buena parte del texto y en el supuesto de que éste les venga resultando interesante.

Hace unos días, hablando con un colega, al decir yo que el libro empezaba analizando los «sistemas de ecuaciones lineales», en lugar de hacerlo con los «espacios vectoriales», él torció el gesto y mostró su desagrado. Me dijo que las cosas deben empezar por el principio, que los «espacios vectoriales» son, en el álgebra lineal, el principio, el principio que manda la lógica y que, en la matemática, es la lógica la que ha de mandar. Mi punto de vista es otro: al que se inicia en el estudio del «álgebra lineal» (o de cualquier otra parte de las matemáticas), hay que presentarle las cosas, al menos en los comienzos, de manera que él llegue a participar, de algún modo, en su elaboración, que pueda apreciar cómo se ha ido configurando lo que se le explica, que llegue a ver la génesis y la conveniencia de los conceptos que van apareciendo ante sus ojos.

Obviamente, al concluir el curso, las cosas han de quedar lógicamente estructuradas; lo uno no impide lo otro. Una cosa es la meta a la que nos dirigimos y otra el camino que se utiliza para ello; en la meta gobierna la lógica, el camino lo marca la didáctica. Si, el que estudia una materia, llega a percatarse del proceso que, desde las situaciones fáciles y concretas, conduce a los estados finales, abstractos y complejos, podrá estar él, entonces, en condiciones de hacer otro tanto cuando le toque hacerlo.

Otro de los palillos que conviene tocar es lo que los físicos llaman «los escalares», que son cimiento del álgebra lineal, que forman su «cuerpo base». Aun cuando a alguien pueda parecerle extraño, para nosotros, en este curso de álgebra lineal, el que los escalares sean unos u otros tiene escasa importancia. Si a ello le unimos que, al menos para el principiante, las dificultades propias del álgebra lineal no son despreciables, hemos decidido no complicarnos demasiado la vida con los escalares: el mayor protagonismo corre a cargo de los números reales; en los casos que les son propios, se utilizan los números complejos; cuando, en un caso general, nos hemos sentido obligados a considerar escalares cualesquiera, se aconseja al lector que, a la menor dificultad con ellos, sustituya «el cuerpo K de los escalares» por «el cuerpo $\mathbb{R}$ de los números reales».

Un tratado sobre «álgebra lineal», ¿ha de ocuparse de los conjuntos, las relaciones binarias, las aplicaciones, los grupos, los anillos y los cuerpos? Estas materias, que constituyen lo que llamaremos «álgebra básica», deben ser conocidas, aunque sólo sea sucintamente, como un escalón previo al «álgebra lineal». Ahora bien, dado que en los estudios preuniversitarios, de los que proceden nuestros posibles lectores, no se suele abordar el «álgebra básica» con la amplitud que, para nosotros, sería deseable, nos hemos visto obligados a no silenciar dichas cuestiones. Les hemos dado un tratamiento especial: se las incluye como apéndices, al final del texto, tratándolas con la debida extensión, sin omitir demostraciones, pero sin ocuparnos de sus aspectos prácticos.

Algo que ha de llamar la atención del lector, que al autor le parece extravagante, pero que, al decir del editor, ha de resultar práctico y provechoso, es el artificio de oscurecer o sombrear el papel allí donde van escritas las partes del texto que se desean destacar. Se ha «sombreado» lo más esencial: las definiciones (las que son principales) y los teoremas y resultados más importantes. Este asunto del «sombrear» no ha conducido a un excesivo oscurecimiento del libro, pues éste está estructurado de modo que sus puntos fundamentales, definiciones y resultados básicos, se presentan compendiados y libres de aditamentos; aparecen desglosados de todos los comentarios, complementos, observaciones o generalizaciones, que se hagan al respecto. Junto a cada uno de dichos textos sombreados, tras de ellos y formando cuerpo aparte, también se incluyen los pertinentes ejemplos y ejercicios, que son excelente recurso para clarificar los conceptos y afianzar las ideas.

Aunque sea poco, algo hemos de decir sobre si le conviene o no al lector, que aquí se le supone principiante, comprobar todas y cada una de las afirmaciones que se le hacen. Nosotros no hemos omitido ninguna demostración, pues creemos que no se le deben ocultar al que estudia; ello no supone que, a nuestro parecer, deban ser todas obligatorias. Cada uno procederá según mejor le convenga, pero un criterio que suele dar buenos resultados es el de omitir pocas demostraciones, omitir sólo aquellas, excesivamente laboriosas o que, por arti-

ficiales, no nos puedan servir de guía cuando debamos abordar, por nuestra cuenta, los problemas que se nos puedan presentar.

Para acabar, vayamos al apartado mas grato de escribir: el de los agradecimientos. En primer lugar he de mostrar mi gratitud a los colegas que, en mayor o menor medida, me han apuntado ideas y sugerencias valiosas que, obviamente, he aprovechado; muestro aquí mi reconocimiento, a José Manuel Vega, a Ignacio Parra, a José Antonio Nicolás, a Manuel Martínez, a Damián Rivas, a Eduardo Ahedo, a Julio Ramírez, a Carlos Vázquez, a José María Nogueira, a Francisco Mancebo, a Carlos Alvarez, a Victoria Lapuerta. Cito aquí, gustoso, a Eva Villacieros, agradeciéndole su gran profesionalidad en la difícil tarea de escribir textos matemáticos. Es de justicia señalar aquí el esmero con el que la editorial McGraw-Hill Interamericana de España ha emprendido la edición de este libro; mi gratitud a Antonio García-Maroto, su director, y a Juan Stumpf, editor. En esto del agradecer, he de hacer mención especial a Carmina, mi mujer, y a mis hijos Carmen y Juan, aunque sólo fuera por la disposición desinteresada y animosa con la que han respondido a las muchas horas que, dedicado al libro, les hurté a ellos.

Verano de 1993, Año Jacobeo
Santiago de Compostela

PARTE I

INICIACIÓN: LINEALIDAD Y RANGO

PRELUDIO *introito o exordio, en el que se habla algo, por lo general poco y no siempre bastante, acerca de los «sistemas de ecuaciones lineales», el «rango», las «matrices» y los «determinantes», queriendo así ofrecer al lector, tanto una herramienta que le permita manejarse con soltura y eficacia, en lo que luego vendrá, como una referencia y un cobijo a los que poder acogerse cuando, pronto, se estudien conceptos más abstractos y generales.*

Introducción.
1. Sistemas de ecuaciones lineales: el método de Gauss.
2. Rango (de vectores y de matrices).
3. Operaciones con matrices: matriz inversa.
4. Determinantes.

Ejercicios y problemas.

Introducción

Este modo de iniciar un curso de «Algebra Lineal», hablando de los «sistemas de ecuaciones lineales» y del «rango», no es ninguna tontería. Hay importantes motivos para ello: Bien es verdad que, desde un punto de vista lógico y formal, quizá habría que pensar en emprender la tarea presentando en primer lugar a los «espacios vectoriales», luego a los «homomorfismos» o «aplicaciones lineales» entre espacios vectoriales, más tarde a las «matrices» y, con todo ello y algo más, abordar debidamente el estudio de los «sistemas de ecuaciones lineales» y de otras cuestiones. Creemos, no obstante, que la didáctica dice otra cosa.

Los conceptos deberán presentarse de la manera que mejor puedan entenderse, la cual no es, generalmente, la que demanda la lógica. Ello no debe impedir que, al final del proceso de aprendizaje, los conocimientos adquiridos terminen formando el entramado adecuado, de suerte que, a la postre, se consiga que los saberes queden pertinentemente estructurados.

Con el correr de los años, han sido muchos los problemas que han desembocado en sistemas de ecuaciones lineales, de las que ha habido que hallar las soluciones. Al abrigo de estos procesos de búsqueda de soluciones, y al de otros desarrollos similares, se han ido perfilando los conceptos de vector, matriz, rango, homomorfismo, etc.; no ha ocurrido al revés. Por ello, entendemos que, al menos ahora, al comienzo, todos estos conceptos se deben dar a conocer respetando su «antigüedad». Piénsese que aquello que primero se resolvió fue, a no dudarlo, lo que era más fácil de resolver.

Es obvio que con el poco bagaje de conocimientos que, a estas alturas, se nos supone, no nos va a ser posible llegar al fondo de las cuestiones que aquí se abordan; tiempo tendremos para ello más adelante.

Este capítulo introductorio pretende cubrir varios objetivos, todos ellos valiosos. Este capítulo es el escalón que permitirá acceder a una estancia que está situada más alta de lo que es ordinario. Este capítulo es la base en la que sustentar los conceptos abstractos que luego vendrán. Este capítulo será la herramienta o instrumento que permitirá manejarse con soltura en lo que sigue. Este capítulo es el germen de multitud de ideas y conceptos que pueblan el «álgebra lineal».

CAPITULO

1

Sistemas de ecuaciones lineales; el método de Gauss

α PRIMERAS DEFINICIONES. EQUIVALENCIA

Según ya reza en el título de este libro, los problemas que nos van a ocupar son los «lineales», entendiendo por tales a aquellos en los que al duplicar, triplicar, etc., la causa, acontece que se duplica, triplica, etc., el efecto; también se requiere que, al sumar las causas, se sumen los efectos. Así, por ejemplo, si las causas vienen representadas por x e y, que denotan a dos números cualesquiera, y el efecto es $\alpha = 3x - 2y$, este α depende linealmente de los x e y.

Los problemas directos son aquellos en los que los datos conducen directamente, sin rodeos, a los resultados. Si lo que se conocen son los resultados y se buscan los datos que se necesitarían para obtener dichos resultados, se está considerando un problema inverso (o recíproco). Los sistemas de ecuaciones lineales tienen esta condición de problema inverso; en ellos, hay que hallar los datos (x e y, por ejemplo) que, en un problema de tipo lineal, conducirían a unos resultados que nos son conocidos (α y β, por ejemplo). Así, en un sistema como él

$$\left.\begin{aligned} 3x - 2y &= \alpha \\ 4x + 5y &= \beta \end{aligned}\right\}$$

hay que hallar los valores de x e y (que se llaman incógnitas) que conducen a los resultados α y β (que son dados).

¡CUIDADO CON LA PALABRA «LINEAL»!

Hay que avisar que, en el hablar ordinario de las gentes, hoy día se está utilizando la palabra lineal para denotar a lo que es constante. Así, cuando se dice que se ha producido un «aumento lineal del sueldo» de unos trabajadores y que dicho aumento ha sido de P pesetas, se nos quiere indicar que a todos y cada uno de los trabajadores se le ha aumentado el sueldo en una misma cantidad, P pesetas, constante para todos. Si en esto de las subidas de sueldos, la linealidad significase lo mismo que en Matemáticas, una subida lineal de un 10 por 100, pongamos por caso, significaría que al que gana 100 se le aumenta en 10 y al que gana 200 se le aumenta en 20; esta subida no es «constante» es «proporcional» a lo que se gana.

1.1. SISTEMAS DE ECUACIONES LINEALES

Debemos empezar precisando el significado de los distintos términos que vamos a utilizar en lo que sigue:

1

DEFINICIONES

[001]

Un sistema de m ecuaciones lineales en las n incógnitas $x_1, x_2, ..., x_n$ es todo conjunto de relaciones del tipo:

$$\left.\begin{array}{ll} \text{Ecuación primera:} & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \text{Ecuación segunda:} & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \text{Ecuación } m\text{-ésima:} & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\} \quad [S]$$

Los a_{ij} (coeficientes) y los b_i (términos independientes) son escalares dados; es decir, elementos de un cuerpo cualquiera(*). Se dice que los escalares $\alpha_1, \alpha_2, ..., \alpha_n$ constituyen una *solución* de S si al tomar $x_1 = \alpha_1$, $x_2 = \alpha_2, ..., x_n = \alpha_n$, las m ecuaciones se convierten en igualdades; *resolver* S es hallar todas sus soluciones. Se dice que un sistema es *compatible* si tiene alguna solución y que es *incompatible* si carece de ellas; cuando es compatible, se dice que es *determinado* si tiene una sola solución e *indeterminado* si tiene más de una.

(*) El cuerpo K de los escalares puede ser cualquiera. Puede suponerse, si ello facilita la comprensión de estas cuestiones, que K es el cuerpo real ($K = \mathbb{R}$) o que es el cuerpo complejo ($K = \mathbb{C}$).

Discutir un sistema de ecuaciones es analizar si éste tiene soluciones y, caso de tenerlas, cuántas tiene.

EJEMPLOS

1. $\left.\begin{array}{l} x+y+z=6 \\ x+y-z=0 \\ x\quad\ +z=4 \\ x+y\quad\ =3 \end{array}\right\}$ Este sistema (en las incógnitas x, y, z) tiene solución única que es la $(x, y, z) = (1, 2, 3)$; se trata, pues, de un sistema compatible y determinado.

2. $\left.\begin{array}{l} x+y+\ z=6 \\ x+y-\ z=0 \\ x+y\quad\ =3 \\ x+y+2z=3 \end{array}\right\}$ Este sistema es compatible indeterminado, ya que tiene infinitas soluciones; éstas son todas las $(x, y, z) = (\alpha, 3-\alpha, 3)$, donde α es un número real cualquiera.

3. $\left.\begin{array}{l} x+\ y+z=6 \\ x+\ y-z=0 \\ -2x-2y\quad\ =1 \end{array}\right\}$ Este sistema es incompatible, carece de soluciones. Nótese que al sumar las tres ecuaciones se llega a $0 = 7$.

SIGNO SUMATORIO; INDICES MUDOS E INDICES LIBRES

1. Para escribir de manera abreviada una suma como la $P_1 + P_2 + \cdots + P_n$, se recurrirá a la letra griega Σ (sigma mayúscula), que recibe el nombre de *signo sumatorio*, y se pone:

$$P_1 + P_2 + \cdots + P_n = \sum_{j=1}^{n} P_j$$

Aquí se ha recurrido a la letra j, que se llama *índice mudo*, que no tiene un valor determinado; se supone que j varía recorriendo todos los valores naturales 1, 2, ..., n. La letra j se puede sustituir por otra letra cualquiera, sin que por ello varíe el resultado; esto es:

$$\sum_{j=1}^{n} P_j = \sum_{k=1}^{n} P_k = \sum_{\alpha=1}^{n} P_\alpha$$

Así, por ejemplo, las m ecuaciones del sistema S (de [001]) pueden ponerse en la forma:

$$\sum_{j=1}^{n} a_{1j}x_j = b_1,\ \sum_{j=1}^{n} a_{2j}x_j = b_2,\ \cdots,\ \sum_{j=1}^{n} a_{mj}x_j = b_m \quad [S]$$

2. Por otra parte, para denotar de manera conjunta a todas y cada una de las m ecuaciones del sistema S, se puede recurrir a otro índice, que llamaremos i y representará a cualquiera de los números 1, 2, ..., m, y poner:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad \text{para } i = 1, 2, ..., m \quad [S]$$

Aquí i es un *índice libre*; para cada valor que se le dé a i, se va a obtener una de las ecuaciones del sistema S.

3. Recurriendo ahora a un índice libre i (que varía de 1 a m) y a un índice mudo (que varía de 1 a n), el sistema S se puede expresar como sigue:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \quad (i = 1.\ 2,\ ...,\ m) \quad [S]$$

EXPRESION MATRICIAL DE UN SISTEMA

Lo que distingue a los sistemas de ecuaciones lineales, diferenciando a unos de otros, son sus coeficientes y sus términos independientes. Para caracterizar a un sistema de ecuaciones, como el

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \quad (i = 1, 2, ..., m) \quad \begin{pmatrix} \text{de } m \text{ ecuaciones: } i = 1, 2, ..., m \\ \text{con } n \text{ incógnitas: } j = 1, 2, ..., n \end{pmatrix}$$

basta con saber los valores que toman los siguientes escalares:

$$\begin{cases} a_{ij} \text{ (coeficiente que afecta a la incógnita } j\text{-ésima de la ecuación } i\text{-ésimas)} \\ b_i \text{ (término independiente de la ecuación } i\text{-ésimas)} \end{cases}$$

Un sistema de ecuaciones queda, pues, determinado en cuanto se conozca la tabla rectangular, o matriz, que forman sus coeficientes y sus términos independientes, situados todos ellos en la misma posición relativa que ocupaban en el sistema. Nótese que esta tabla es lo que queda del sistema al prescindir en él de las incógnitas y de los signos (+ e =). Concretando:

[002]

Sea dado un sistema de m ecuaciones lineales con n incógnitas, al que llamaremos $[S]$:

$$\left.\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ &a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{aligned}\right\} \quad [S]$$

Se llaman *matriz de los coeficientes*, matriz de los términos independientes y *matriz del sistema* (o ampliada) a las siguientes tablas rectangulares(*) A, B y $\mathbb{AB}$ formadas por escalares:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}; \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}; \quad \mathbb{AB} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

De manera abreviada y, por el momento, también simbólicamente, el sistema S se expresará poniendo:

$$AX = B, \quad \text{donde } X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

se llama columna o vector columna de las incógnitas

(*) Posteriormente, se precisará y se completará la definición de matriz. De momento nos basta con saber que la matriz A, por ejemplo, es una tabla rectangular de $m \cdot n$ escalares dispuestos en m filas y n columnas. La fila i-ésima de A es la sucesión de escalares $a_{i1}a_{i2}\cdots a_{in}$; la columna j-ésima de A es la sucesión de escalares $a_{1j}a_{2j}\cdots a_{mj}$ (que en A aparecen dispuestos verticalmente). Se dice que el elemento a_{ij}, que está situado en la fila i y en la columna j, es el elemento que ocupa el lugar (i, j) o ij.

SISTEMAS HOMOGENEOS

Los sistemas de ecuaciones que tienen nulos sus términos independientes, se llaman homogéneos. Esto es:

[003]

Un sistema de ecuaciones lineales se dice homogéneo si es del tipo:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array}\right\} \quad [H]$$

(Abreviadamente, se pondrá $AX = O$, donde O es matriz (de una sola columna) cuyos elementos son todos nulos)

matriz ampliada = matriz coeficientes.

Todo sistema de ecuaciones lineales homogéneo $[H]$ tiene, al menos, la solución $x_1 = x_2 = \cdots = x_n = 0$, llamada solución trivial o nula, por lo que siempre es compatible. Será compatible indeterminado si tiene alguna solución no nula.

Las combinaciones lineales de las soluciones de un sistema homogéneo son, también, soluciones de éste; es decir:

$$\bar{\alpha} \text{ y } \bar{\beta} \text{ son soluciones de } H \Rightarrow$$
$$\Rightarrow \lambda\bar{\alpha} + \mu\bar{\beta} \text{ es solución de } H \ (\lambda \text{ y } \mu \text{ escalares}) \quad (*)$$

(*) Por razones de brevedad, se ha llamado $\bar{\alpha} = (\alpha_1, \ldots, \alpha_n)$, $\bar{\beta} = (\beta_1, \ldots, \beta_n)$ y $\lambda\bar{\alpha} + \mu\bar{\beta} = (\lambda\alpha_1 + \mu\beta_1, \ldots, \lambda\alpha_n + \mu\beta_n)$. $\sum_{j=1}^{n} \alpha_j = \bar{\alpha}$

COMPROBACION

Expresando el sistema homógeneo en la forma abreviada $\Sigma\, a_{ij}x_j = 0$, como $\bar{\alpha}$ y $\bar{\beta}$ son soluciones, se sabe que

$$\sum_{j=1}^{n} a_{ij}\alpha_j = 0 \quad \text{y} \quad \sum_{j=1}^{n} a_{ij}\beta_j = 0 \quad (\text{para } i = 1, 2, \ldots, m)$$

Por tanto, $\lambda\bar{\alpha} + \mu\bar{\beta}$ también es solución, ya que (para $i = 1, 2, \ldots, m$):

$$\sum_{j=1}^{n} a_{ij}(\lambda\alpha_j + \mu\beta_j) = \sum_{j=1}^{n} (\lambda a_{ij}\alpha_j + \mu a_{ij}\beta_j) =$$

$$= \lambda \sum_{j=1}^{n} a_{ij}\alpha_j + \mu \sum_{j=1}^{n} a_{ij}\beta_j = \lambda \cdot 0 + \mu \cdot 0 = 0$$

EJERCICIO

Atendiendo a la «cantidad» de elementos que tienen los conjuntos, a éstos se les clasifica en FINITOS E INFINITOS. Si C es un conjunto finito, se dirá que C tiene un número finito de elementos. Si C es un conjunto finito, se dirá que C tiene infinitos elementos. No es correcto decir, en este último caso, que C tiene un número infinito de elementos; infinito no es un número.

Se desea comprobar que, si un sistema de ecuaciones lineales tiene más de una solución, entonces el sistema tiene infinitas soluciones. Más exactamente, vamos a probar que:

Dado un sistema de ecuaciones lineales

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \quad (\text{con } i = 1, 2. \ldots, m)$$

si

$$\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

y

$$\bar{\beta} = (\beta_1, \beta_2, \ldots, \beta_n)$$

son dos raíces distintas del sistema, entonces se verifica que

a) $\bar{\alpha} - \bar{\beta} = (\alpha_1 - \beta_1, \alpha_2 - \beta_2, \ldots, \alpha_n - \beta_n)$ es raíz del sistema homogéneo $\sum_j a_{ij}x_j = 0$.

b) $\bar{\alpha} + \lambda(\bar{\alpha} - \bar{\beta})$ es raíz del sistema dado $\sum_j a_{ij}x_j = b_i$ para todo escalar λ.

RESOLUCION

a) $\sum_j a_{ij}(\alpha_j - \beta_j) = \sum_j a_{ij}\alpha_j - \sum_j a_{ij}\beta_j = b_i - b_i = 0$

luego $\alpha - \bar{\beta}$ es raíz del sistema homogéneo.

b) $\sum_j a_{ij}[\alpha_i + \lambda(\alpha_i - \beta_i)] = \sum_j a_{ij}\alpha_i + \lambda \sum_j a_{ij}(\alpha_j - \beta_j) = b_i + \lambda \cdot 0 = b_i$

luego $\bar{\alpha} + \lambda(\bar{\alpha} - \bar{\beta})$ es raíz del sistema dado para cualquier λ.

1.2. SISTEMAS EQUIVALENTES

Para resolver un sistema de ecuaciones, se le irá «reduciendo» a otros sistemas que sean más simples; esto es, se buscarán sistemas que tengan las misma soluciones que el dado y que resulten cada vez más fáciles de resolver. A los sistemas que tienen las mismas soluciones se les llama equivalentes; de ellos nos ocupamos ahora:

EQUIVALENCIA DE SISTEMAS; OPERACIONES ELEMENTALES

[004]

Dos sistemas de ecuaciones (con las mismas incógnitas) se dicen *equivalentes* si tienen las mismas soluciones, o sea, si toda solución de cada uno de ellos es también solución del otro.

Propiedad fundamental de la equivalencia.–Un sistema de ecuaciones lineales es equivalente a cualquiera de los sistemas que resultan de realizar, en aquél, *operación elemental*. Se llaman operaciones elementales a las siguientes manipulaciones:

1.° Intercambiar el orden con el que figuran las ecuaciones en el sistema.
2.° Multiplicar una de las ecuaciones por cualquier escalar no nulo.
3.° Sumarle, a una de las ecuaciones, otra cualquiera de ellas.
4.° Aplicar, reiteradamente, cualesquiera de las operaciones anteriores (en particular, sumarle a una ecuación, cualquier combinación lineal de las demás)(*).

NOTA: Hay quienes consideran también como operación elemental (cosa que no haremos nosotros) a la siguiente: suprimir una ecuación que sea combinación lineal de otras ecuaciones.

(*) Por definición, las combinaciones lineales de las ecuaciones $a_1x + b_1y + c_1z = d_1$ y $a_2x + b_2y + c_2z = d_2$ son las ecuaciones

$$(\alpha a_1 + \beta a_2)x + (\alpha b_1 + \beta b_2)y + (\alpha c_1 + \beta c_2)z = \alpha d_1 + \beta d_2$$

donde α y β son escalares cualesquiera.

Nótese que entre las operaciones elementales del apartado 4.°, cabe destacar la siguiente: sustituir una cualquiera de las ecuaciones por una combinación lineal de otras y de ella misma de manera que el coeficiente que multiplica a ésta sea no nulo.

Vamos aquí a simplificar la notación llamando

$$e_i(\bar{x}) \equiv a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - b_i \quad (\text{para } i = 1, 2, \ldots, m)$$

con lo que el anterior sistema de ecuaciones S (véase [002]) se expresará en la forma S'

$$\left.\begin{array}{l} e_1(\bar{x}) = 0 \\ e_2(\bar{x}) = 0 \\ \cdots\cdots\cdots\cdots \\ e_m(\bar{x}) = 0 \end{array}\right\} \quad [S']$$

Con este tipo de notación, la propiedad anterior asegura que, por ejemplo, son equivalentes los siguientes sistemas:

$$\left.\begin{aligned} e_1(\bar{x}) &= 0 \\ e_2(\bar{x}) &= 0 \\ e_3(\bar{x}) &= 0 \end{aligned}\right\} \qquad \left.\begin{aligned} 2e_1(\bar{x}) &= 0 \\ 3e_2(\bar{x}) - e_3(\bar{x}) &= 0 \\ e_3(\bar{x}) &= 0 \end{aligned}\right\} \qquad \left.\begin{aligned} e_1(\bar{x}) &= 0 \\ 2e_2(\bar{x}) + 3e_1(\bar{x}) &= 0 \\ \tfrac{1}{2}e_3(\bar{x}) - e_1(\bar{x}) + 5e_2(\bar{x}) &= 0 \end{aligned}\right\}$$

DEMOSTRACION

Los casos 1.º) y 2.º) son obvios; el caso 4.º) se reduce a aplicar reiteradamente los resultados de los casos que le preceden. Por ello, sólo consideraremos el caso 3.º).

Para señalar que $x_1 = \alpha_1,\ x_2 = \alpha_2,\ \ldots,\ x_n = \alpha_n$ es una solución del sistema, se dirá que la solución es la $\bar{x} = \bar{\alpha}$, esto es, se ha llamado

$$\bar{x} = (x_1, x_2, \ldots, x_n) \quad \text{y} \quad \bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

Para verificar que la propiedad es cierta en el caso 3.º), no hay más que darse cuenta de que las siguientes afirmaciones son equivalentes entre sí:

$\bar{\alpha}$ es solución del sistema	se verifican las igualdades	se verifican las igualdades	$\bar{\alpha}$ es solución del sistema
$e_1(\bar{x}) = 0$	$e_1(\bar{\alpha}) = 0$	$e_1(\bar{\alpha}) + e_2(\bar{\alpha}) = 0$	$e_1(\bar{x}) + e_2(\bar{x}) = 0$
$e_2(\bar{x}) = 0$	$e_2(\bar{\alpha}) = 0$	$e_2(\bar{\alpha}) = 0$	$e_2(\bar{x}) = 0$
.....	.....	.....	.....
$e_m(\bar{x}) = 0$	$e_m(\bar{\alpha}) = 0$	$e_m(\bar{\alpha}) = 0$	$e_m(\bar{x}) = 0$

VECTORES

Aquí, para denotar a las sucesiones $(x_1, x_2, \ldots, x_n)$ y $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, de n escalares, se ha recurrido a escribir $\bar{x}$ y $\bar{\alpha}$. A estos elementos se les llamará *vectores*. Un poco más adelante se estudiarán con detenimiento estos objetos matemáticos; de momento, sólo hemos necesitado de sus *combinaciones lineales:* si

$$\bar{\alpha} = (\alpha_1, \ldots, \alpha_n)$$

$$\bar{\beta} = (\beta_1, \ldots, \beta_n)$$

se llama $\lambda\bar{\alpha} + \mu\bar{\beta} = a$

$$(\lambda\alpha_1 + \mu\beta_1, \ldots, \lambda\alpha_n + \mu\beta_n)$$

(λ y μ son escalares).

OBSERVACION

Sea S un sistema de m ecuaciones lineales y sea S' un sistema formado por m combinaciones lineales de aquellas. Es evidente que toda solución de S es, también, solución de S'. Sin embargo, en general puede haber soluciones de S' que no lo sean de S. Así puede ocurrir en los dos ejemplos siguientes, en los que se toma para S' uno de los S'_1 o S'_2:

$$\left.\begin{aligned} e_1(\bar{x}) &= 0 \\ e_2(\bar{x}) &= 0 \\ e_3(\bar{x}) &= 0 \end{aligned}\right\} \ [S] \quad ; \quad \left.\begin{aligned} e_1(\bar{x}) &= 0 \\ e_3(\bar{x}) &= 0 \\ e_1(\bar{x}) + e_3(\bar{x}) &= 0 \end{aligned}\right\} \ [S'_1] \quad ; \quad \left.\begin{aligned} e_1(\bar{x}) - e_2(\bar{x}) &= 0 \\ e_2(\bar{x}) - e_3(\bar{x}) &= 0 \\ e_3(\bar{x}) - e_1(\bar{x}) &= 0 \end{aligned}\right\} \ [S'_2]$$

Tómese, en concreto

$$e_1(\bar{x}) = x_1 + x_2 - 2,\ e_2(\bar{x}) = x_2 + x_3 - 2 \quad \text{y} \quad e_3(\bar{x}) = x_3 + x_1 - 2$$

En este caso, S tiene una única solución que es la $x_1 = 1$, $x_2 = 1$, $x_3 = 1$. El sistema S_1' tiene infinitas soluciones, que son las $x_1 = \lambda$ (cualquiera), $x_2 = 2 - \lambda$ y $x_3 = 2 - \lambda$. El sistema S_2' tiene infinitas soluciones, que son las $x_1 = x_2 = x_3 = \lambda$ (cualquiera).

EJEMPLO

Recurriendo a la propiedad anterior y aplicándola atinadamente, se pueden resolver los sistemas de ecuaciones. Así, por ejemplo, el siguiente sistema S es equivalente a los S' y S'' y la resolución de este último es elemental:

$$\left.\begin{array}{ll} (1.^a) \rightarrow & x + y + 2z = 9 \\ (2.^a) \rightarrow & 3x + 6y - 5z = 0 \\ (3.^a) \rightarrow & 2x + 4y - 3z = 1 \end{array}\right\} \quad [S]$$

$$\left.\begin{array}{ll} (1') = (1.^a) & \rightarrow \quad x + y + 2z = 9 \\ (2') = (2.^a) - 3(1.^a) & \rightarrow \quad 3y - 11z = -27 \\ (3') = (3.^a) - 2(1.^a) & \rightarrow \quad 2y - 7z = -17 \end{array}\right\} \quad [S']$$

$$\left.\begin{array}{ll} (1'') = (1') & \rightarrow \quad x + y + 2z = 9 \\ (2'') = (2') & \rightarrow \quad 3y - 11z = -27 \\ (3'') = 3(3') - 2(2') & \rightarrow \quad z = 3 \end{array}\right\} \quad [S'']$$

La solución de S'' es, evidentemente, la $z = 3$, $y = 2$, $x = 1$, que es también la solución del sistema dado.

ENFOQUE MATRICIAL

Si se conoce la matriz de un sistema, éste también se conoce y, por ello, se podrá resolver. Al ser mucho más cómodo manejar solamente la matriz, en lugar de todo el sistema, así se hará, con lo que simplificaremos el proceso de resolución. En este sentido, es útil observar que la anterior propiedad [004], sobre sistemas equivalentes, puede formularse en los siguientes términos:

[005]

Un sistema de ecuaciones lineales $AX = B$ es equivalente a cualquiera de los que se obtienen de realizar operaciones elementales en las filas de su matriz $\mathcal{AB}$; se llaman operaciones elementales a las siguientes manipulaciones:

1.° Cambiar el orden en el que, en $\mathcal{AB}$, aparecen las filas.
2.° Multiplicar una fila por un escalar no nulo.
3.° Sumarle, a una fila, otra cualquiera de ellas.
4.° Aplicar, reiteradamente, cualesquiera de las operaciones anteriores (en particular, sumarle, a una fila, cualquier combinación lineal de las demás).

EJEMPLO

Para resolver el siguiente sistema S (que es el ya considerado en el último ejemplo), si manejamos (en lugar de todo el sistema) solamente su matriz $\mathcal{AB}$, al realizar operaciones elementales en sus filas, se obtiene

$$\left.\begin{aligned} x+\ y+2z&=9\\ 3x+6y-5z&=0\\ 2x+4y-3z&=1 \end{aligned}\right\}\ [S]\ ;\ \mathcal{AB}=\left[\begin{array}{ccc|c} 1 & 1 & 2 & 9\\ 3 & 6 & -5 & 0\\ 2 & 4 & -3 & 1 \end{array}\right]\begin{array}{l}\leftarrow 1.^{a}\\ \leftarrow 2.^{a}\\ \leftarrow 3.^{a}\end{array}$$

$$\begin{array}{ll} (1')=(1.^{a}) & \rightarrow \\ (2')=(2.^{a})-3(1.^{a}) & \rightarrow \\ (3')=(3.^{a})-2(1.^{a}) & \rightarrow \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 2 & 9\\ 0 & 3 & -11 & -27\\ 0 & 2 & -7 & -17 \end{array}\right]$$

$$\begin{array}{ll} (1'')=(1') & \rightarrow \\ (2'')=(2') & \rightarrow \\ (3'')=3(3')-2(2') & \rightarrow \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 2 & 9\\ 0 & 3 & -11 & -27\\ 0 & 0 & 1 & 3 \end{array}\right]$$

Esta última matriz lo es de un sistema que se resuelve trivialmente; su solución, que es la misma que la de S, es la $x = 1$, $y = 2$, $z = 3$.

β EL METODO DE GAUSS

1.3. MATRICES ESCALONADAS Y SISTEMAS ESCALONADOS

Nos vamos a referir aquí a unos sistemas de ecuaciones, especialmente fáciles de resolver, a los que luego reduciremos todos los demás sistemas: todo sistema de ecuaciones lineales resultará ser equivalente a alguno de los sistemas escalonados.

[006]

Se llaman *matrices escalonadas* a aquellas matrices en las que:

1.° Cada una de las filas (a partir de la 2.ª) comienza con una sucesión de ceros que tiene algún cero más (uno como mínimo) que la fila anterior. Nótese que si hubiera filas formadas sólo por ceros, éstas serían las últimas filas de la matriz.
2.° En las filas en las que algún elemento no es nulo, el primero de éstos (que llamaremos cabecera de la fila) es un 1(*).

Se llaman *matrices escalonadas reducidas* a las matrices escalonadas en las que, además, se verifica:

3.° En las columnas en las que están ubicadas las cabeceras de las filas, todos los demás elementos son nulos.

A la matriz traspuesta(**) de una matriz escalonada se la suele llamar «matriz escalonada por columnas».

Un sistema de ecuaciones lineales se dice que es un *sistema escalonado* (respectivamente, *sistema escalonado reducido*) si su matriz de coeficientes es escalonada (respectivamente, escalonada reducida).

(*) Este requisito podría omitirse sin que, por ello, variase, en esencia, el proceso y los razonamientos que haremos a continuación.

(**) La matriz traspuesta de A se obtiene al cambiar, en A, las filas por las columnas.

MATRIZ ESCALONADA

$$\begin{bmatrix} 1 & * & * & * & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

0 = elementos nulos
1 = elementos unidad
* = elementos cualesquiera (todos los que están sobre cada uno de los elementos 1 pueden hacerse nulos)

EJEMPLOS

En cada uno de los siguientes ejemplos, con los que se pretenden aclarar las anteriores definiciones, se da un sistema de ecuaciones escalonadas, su matriz ampliada y su solución, si la tiene:

SISTEMA:	MATRIZ:	SOLUCION:
$\left.\begin{array}{rrrrr} x & +3z & -v & = & 2 \\ & y+5z & & = & 7 \\ & & u+2v & = & -1 \end{array}\right\}$	$\left[\begin{array}{ccccc\|c} 1 & 0 & 3 & 0 & -1 & 2 \\ 0 & 1 & 5 & 0 & 0 & 7 \\ 0 & 0 & 0 & 1 & 2 & -1 \end{array}\right]$	$\begin{cases} x = 2 - 3z + v \\ y = 7 - 5z \\ z = \text{cualquiera} \\ u = -1 - 2v \\ v = \text{cualquiera} \end{cases}$
$\left.\begin{array}{rrrrr} x & +3z & -v & = & 2 \\ & y+5z & & = & 7 \\ & & u+2v & = & -1 \\ & & & 0 = & 0^{(**)} \end{array}\right\}$	$\left[\begin{array}{ccccc\|c} 1 & 0 & 3 & 0 & -1 & 2 \\ 0 & 1 & 5 & 0 & 0 & 7 \\ 0 & 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$	$\begin{cases} x = 2 - 3z + v \\ y = 7 - 5z \\ z = \text{cualquiera} \\ u = -1 - 2v \\ v = \text{cualquiera} \end{cases}$
$\left.\begin{array}{rrrrr} x & +3z & -v & = & 2 \\ & y+5z & & = & 7 \\ & & u+2v & = & -1 \\ & & & 0 = & 4^{(**)} \end{array}\right\}$	$\left[\begin{array}{ccccc\|c} 1 & 0 & 3 & 0 & -1 & 2 \\ 0 & 1 & 5 & 0 & 0 & 7 \\ 0 & 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{array}\right]$	sistema incompatible
$\left.\begin{array}{rrrrr} x+4y & +z & +u & = & 1 \\ & y+2z & -u & = & 4 \\ & & z-4u & = & 9 \\ & & u & = & -2 \end{array}\right\}$	$\left[\begin{array}{cccc\|c} 1 & 4 & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 & 4 \\ 0 & 0 & 1 & -4 & 9 \\ 0 & 0 & 0 & 1 & -2 \end{array}\right]$	$\begin{cases} u = -2 \\ z = 1 \\ y = 0 \\ x = 2 \end{cases}$

DISCUSION DE LOS SISTEMAS ESCALONADOS

[007]

Sea E un sistema escalonado de m ecuaciones lineales con n incógnitas; sea M la matriz de los coeficientes de E. De las m filas de M, llamaremos r al número de ellas que tienen algún elemento no nulo; las $m - r$ últimas filas de M sólo contienen ceros. Respecto de la compatibilidad del sistema E, se verifica:

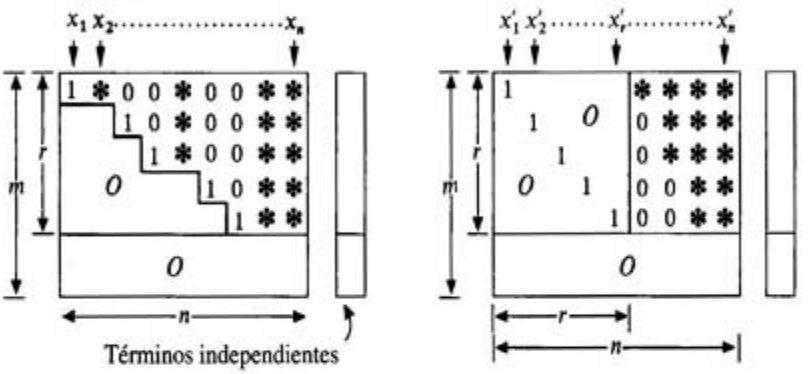
1.° E es compatible si y sólo si sus últimos $m - r$ términos independientes son todos nulos.

2.° Suponiendo que los $m - r$ últimos términos independientes son nulos, el sistema es compatible determinado si $r = n$ y es compatible indeterminado si $r < n$ (nótese que no es posible que $r > n$).

(**) Estos tipos de ecuaciones no son usuales y, ello, sobre todo cuando, como aquí, son datos de partida. Sin embargo, pueden aparecer, y de hecho aparecen con frecuencia, cuando el sistema de ecuaciones es el resultado de aplicar operaciones elementales en un sistema dado.

COMPROBACION

Esta verificación puede reducirse al caso de sistemas escalonados reducidos. En efecto: si E no es reducido, recurriendo a ciertas operaciones elementales, que se localizan trivialmente, se obtiene un sistema E' equivalente a E, escalonado reducido, con el mismo valor de r y con los mismos $m-r$ últimos términos independientes. Si para E' las cosas son como se dice en el enunciado, es obvio, entonces, que también lo son para E y, por ello, bastará considerar aquí a los sistemas escalonados reducidos.



Supongamos, pues, que E es un sistema escalonado reducido. La discusión de E resultará más fácil de obtener si, previamente, alteramos el orden de las incógnitas $x_1, x_2, \ldots, x_n$, que pasaremos a llamarlas $x'_1, x'_2, \ldots, x'_n$, y ello del modo que ahora se indica: si las cabeceras de las sucesivas filas de la matriz M son los elementos de lugares $(1, 1), (2, h), (3, k), \ldots, (r, l)$, entonces tomamos

$$x'_1 = x_1 \quad x'_2 = x_h \quad x'_3 = x_k \quad \cdots \quad x'_r = x_l$$

las restantes $m-r$ incógnitas se tomarán en el mismo orden relativo que tienen. Haciendo esto, el sistema E queda de la forma siguiente (para ciertas constantes: coeficientes c_{ij} y términos independientes h_i):

$$\left.\begin{aligned}
x'_1 &= h_1 - (c_{1r+1}\ x'_{r+1} + \cdots + c_{1n}\ x'_n)\\
x'_2 &= h_2 - (c_{2r+1}\ x'_{r+1} + \cdots + c_{2n}\ x'_n)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
x'_r &= h_r - (c_{rr+1}\ x'_{r+1} + \cdots + c_{rn}\ x'_n)\\
0 &= h_{r+1}\\
&\cdots\cdots\\
&\cdots\cdots\\
0 &= h_m
\end{aligned}\right\} \quad [E_1]$$

Aquí, los h_i son los términos independientes de E; los c_{ij} son unos ciertos escalares, que forman parte de la matriz M. A la vista de la forma E_1 de presentar el sistema E, resultan evidentes las afirmaciones 1.ª y 2.ª del anterior enunciado.

1.4. EL METODO DE GAUSS

Todo sistema de ecuaciones lineales S se va a poder reducir a uno escalonado E; es decir, realizando adecuadas operaciones elementales en S, siempre se conseguirá un sistema escalonado E, equivalente a S. Si el sistema S es compatible, como la resolución de E es trivial, según acabamos de ver en el apartado anterior, resulta que tenemos ya a nuestro alcance un procedimiento práctico para resolver sistemas de ecuaciones lineales, que no es otro, en esencia, que el método de eliminación de Gauss, que aquí se considera.

ELIMINACION SUCESIVA DE INCOGNITAS

[008]

Dado un sistema S de m ecuaciones en las n incógnitas $x_1, x_2, ..., x_n$, eliminar la incógnita x_1 es obtener un sistema S' equivalente a S, también de m ecuaciones y con las mismas n incógnitas, tal que la incógnita x_1 no interviene en $n-1$ de las ecuaciones de S, es decir, del siguiente tipo:

$$\left.\begin{array}{l} e_1(x_1, x_2, ..., x_n) = 0 \\ e_2(x_1, x_2, ..., x_n) = 0 \\ \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ e_m(x_1, x_2, ..., x_n) = 0 \end{array}\right\} [S]; \qquad \left.\begin{array}{r} f_1(x_1, x_2, ..., x_n) = 0 \\ f_2(x_2, ..., x_n) = 0 \\ \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ f_m(x_2, ..., x_n) = 0 \end{array}\right\} [S']$$

OBSERVACIONES

Supongamos que se ha eliminado x_1 y llamemos S_1 al sistema de las $m-1$ últimas ecuaciones de S_1 (que tiene $n-1$ incógnitas). Si en S_1 pudiéramos eliminar ahora otra incógnita (la x_2) y seguir así sucesivamente, llegaríamos, al final del proceso, a un último sistema S^u, *equivalente a S*, en el que cada ecuación tendría una incógnita menos que la ecuación anterior; S^u sería del tipo:

$$\left.\begin{array}{r} g_1(x_1, x_2, x_3, ..., x_n) = 0 \\ g_2(x_2, x_3, ..., x_n) = 0 \\ g_3(x_3, ..., x_n) = 0 \\ \cdots\cdots\cdots\cdots \\ g_m(?) = 0 \end{array}\right\} [S^u]$$

Con el signo de interrogación que se ha puesto en la última ecuación de S^u, se quiere señalar que este proceso puede concluir de manera distinta, dependiendo de las ecuaciones que se consideren y, en particular, de si m es mayor, menor o igual que n. Este sistema S^u es, previsiblemente, más fácil de resolver que el S. En el caso de compatibilidad, se comenzará resolviendo la última ecuación de S^u, que nos dará el valor de x_m, si sólo tiene esta incógnita, o una cierta x_h en función de las $x_{h+1}, x_{h+2}, \ldots, x_m$. Avanzando ahora de abajo a arriba y llevando a cada ecuación el resultado obtenido en la siguiente, se van despejando las sucesivas incógnitas, hasta llegar a la primera ecuación, de la que se obtiene finalmente x_1.

En los sistemas de ecuaciones lineales, el anterior proceso de eliminación se puede realizar fácilmente recurriendo a las operaciones elementales; nótese que, en esencia, se trata de conseguir un sistema equivalente al dado que sea escalonado. Esto es lo que se hace en el siguiente ejemplo:

EJEMPLO

Consideremos el siguiente sistema de ecuaciones, cuya matriz ampliada es la que se indica a su derecha (en este sistema, a denota a un parámetro):

$$\left.\begin{aligned} x+2y+\quad z &= 1\\ 2x+\ y+\quad 3z &= -4\\ x-\ y+(a+2)z &= -3a-5\\ 4x+2y+(a+6)z &= -3a^2-8 \end{aligned}\right\}\ [S];\quad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & -4 \\ 1 & -1 & a+2 & -3a-5 \\ 4 & 2 & a+6 & -3a^2-8 \end{array}\right]\begin{array}{l}\leftarrow 1.^a\\ \leftarrow 2.^a\\ \leftarrow 3.^a\\ \leftarrow 4.^a\end{array}$$

Al intentar resolver el sistema e ir, para ello, eliminando incógnitas mediante operaciones elementales en las filas de la matriz, obtenemos sucesivamente las siguientes matrices, que conducen al sistema $[E]$ que es equivalente al dado y escalonado:

$$\begin{array}{ll} 1' = 1.^a & \rightarrow \\ 2' = 2.^a - 2(1.^a) & \rightarrow \\ 3' = 3.^a - 1.^a & \rightarrow \\ 4' = 4.^a - 4(1.^a) & \rightarrow \end{array}\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & -3 & 1 & -6 \\ 0 & -3 & 1+a & -3a-6 \\ 0 & -6 & 2+a & -3a^2-12 \end{array}\right]$$

$$\begin{array}{ll} 1'' = 1' & \rightarrow \\ 2'' = -2' & \rightarrow \\ 3'' = 3' - 2' & \rightarrow \\ 4'' = 4' - 2(2') & \rightarrow \end{array}\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 3 & -1 & 6 \\ 0 & 0 & a & -3a \\ 0 & 0 & a & -3a^2 \end{array}\right]$$

$$\begin{array}{ll} 1''' = 1'' & \rightarrow \\ 2''' = 2'' & \rightarrow \\ 3''' = 3'' & \rightarrow \\ 4''' = 4'' - 3'' & \rightarrow \end{array}\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 3 & -1 & 6 \\ 0 & 0 & a & -3a \\ 0 & 0 & 0 & -3a(a-1) \end{array}\right];\quad \left.\begin{aligned} x+2y+z &= 1\\ 3y-z &= 6\\ az &= -3a\\ 0 &= -3a(a-1) \end{aligned}\right\}\ [E]$$

La última ecuación de E sólo se verifica si el parámetro a toma uno de los valores $a = 0$ o $a = 1$. Por ello, concluimos que:

- Para $a \neq 0$ y $a \neq 1$, $[S]$ es incompatible.
- Para $a = 0$, $[S]$ es compatible indeterminado, pues es equivalente al sistema

$$\left.\begin{aligned} x + 2y + z &= 1 \\ 3y - z &= 6 \\ 0 &= 0 \\ 0 &= 0 \end{aligned}\right\}, \text{ que tiene las soluciones: } \begin{cases} z = \lambda \text{ cualquiera} \\ y = 2 + \lambda/3 \\ x = -3 - 5\lambda/3 \end{cases}$$

- Para $a = 1$, $[S]$ es compatible determinado, pues es equivalente al sistema:

$$\left.\begin{aligned} x + 2y + z &= 1 \\ 3y - z &= 6 \\ z &= -3 \end{aligned}\right\}, \text{ que tiene sólo la solución: } \begin{cases} z = -3 \\ y = 1 \\ x = 2 \end{cases}$$

METODO DE ELIMINACION DE GAUSS

[009]

Sea S el sistema de ecuaciones lineales $\sum_j a_{ij}x_j = b_i$ ($i = 1, 2, ..., m$ y $j = 1, 2, ..., n$); sea $\mathcal{AB}$ su matriz ampliada.

Se dice que un elemento $a_{ij} \neq 0$ se utiliza de *pivote* hacia abajo (o hacia arriba o ambas cosas) si en las filas de $\mathcal{AB}$ se realizan aquellas operaciones elementales que transforman a a_{ij} en 1 y convierten en 0 a los demás elementos de la columna j que están por debajo (o por arriba o ambas cosas) del a_{ij}. Todo elemento no nulo a_{ij} de A puede utilizarse como pivote.

Utilizando como pivotes a ciertos elementos de A, elegidos de modo idóneo para realizar adecuadas operaciones elementales en las filas de A, siempre es posible transformar S en un sistema equivalente E que sea escalonado e, incluso, que sea escalonado reducido.

Si es compatible, para resolver el sistema S se puede proceder de dos maneras:

1.º Hallando el citado sistema E escalonado, pero no reducido, despejando la incógnita de cabecera de su última ecuación y, avanzando de abajo a arriba, sustituir el valor hallado para cada incógnita en la ecuación anterior, de la que se despeja una nueva incógnita (método de eliminación de Gauss).

2.º Hallando el sistema E escalonado reducido, de cada una de sus ecuaciones se despeja (directamente) la incógnita que esté en su cabecera (eliminación de Gauss-Jordan).

COMPROBACIONES

—Cerciorémonos primero de que cualquier elemento $a_{ij} \neq 0$ se puede utilizar de pivote. Para ello, realicemos las siguientes operaciones elementales en las filas de $\mathcal{AB}$: Primero dividamos la fila de lugar i por a_{ij}, con lo que el elemento de lugar (i, j) pasa a valer 1; después, a la fila de lugar $h \neq i$ le restamos la fila de lugar i multiplicada por a_{hj}, con lo que el elemento de lugar (h, j) se transforma en 0. Al repetir esta última operación para todos los $h > i$ (ó $h < i$ o ambas cosas), todos los elementos que están por debajo del (i, j) (o por arriba o ambas cosas) pasan a ser 0, con lo que a_{ij} ha podido ser utilizado como pivote.

—Veamos ahora cómo obtener un sistema escalonado E que sea equivalente al S. En la primera columna de A hay algún elemento no nulo(*); intercambiando dos filas de A, si fuera preciso, se obtendría en el lugar (1, 1) un elemento no nulo; utilizándolo de pivote, se transforman en 0 todos los elementos situados por debajo de él. Con esto, la 1.ª columna está ya en la forma deseada. Si todos los elementos que quedan por debajo del (1, 2) son nulos, se deja la 2.ª columna como está; si alguno de ellos es no nulo, intercambiando dos filas si fuera preciso, se le lleva al lugar (2, 2) y, luego, se le utiliza de pivote para anular todos los elementos que están situados por debajo de él. Con esto, las 1.ª y 2.ª columnas están ya en la forma deseada. Se pasa ahora a la 3.ª columna, con la que se procede de manera análoga y se reitera el proceso hasta llegar a la última columna de la matriz de coeficientes, con lo que ésta queda, después de aplicarle las operaciones elementales, en forma escalonada.

—Comprobemos que se puede conseguir que el susodicho sistema escalonado (equivalente a S) sea reducido. Para ello, partimos del sistema escalonado E (no reducido), obtenido como se ha indicado en el párrafo anterior. Fijémonos ahora en la cabecera de cada una de las ecuaciones no nulas de E; a los lugares que ocupan estas cabeceras les llamaremos $(1, 1)$, $(2, i_2)$, $(3, i_3)$, ..., (r, i_r). Utilizando como pivote al elemento de lugar $(1, i_r)$, se convierten en ceros todos los elementos que están por encima de él; luego, utilizando como pivote al elemento de lugar $(r-1, i_{r-1})$, se convierten en ceros todos los elementos situados por encima de él; se reitera este proceso hasta, finalmente, transformar en cero al elemento situado encima del de lugar $(2, i_2)$. De todo ello, resulta que el sistema (equivalente a S) que así se ha obtenido es, obviamente, escalonado reducido.

—Respecto de cómo resolver los sistemas escalonados y los escalonados reducidos, no hay nada importante que añadir a lo ya dicho en su descripción.

OBSERVACIONES (RANGO Y GRADOS DE LIBERTAD)

Sea S un sistema compatible de m ecuaciones lineales con n incógnitas. Supongamos que hemos aplicado el método de Gauss-Jordan y que de él se ha obtenido un sistema escalonado reducido E, equivalente a S.

(*) Si no lo hubiere, en S no aparecería la incógnita x_1, pues todos sus coeficientes serían nulos; esto es, x_2 sería la primera incógnita y no es así.

El sistema E, sin más que alterar adecuadamente el orden de sus incógnitas(*), esto es, de cambiar sus subíndices, se puede expresar en la siguiente forma E_1 (que ya fue analizada al finalizar [007]:

$$E_1 \begin{cases} x_1 = h_1 - (c_{1r+1}x_{r+1} + \cdots + c_{1n}x_n) \\ x_2 = h_2 - (c_{2r+1}x_{r+1} + \cdots + c_{2n}x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_r = h_r - (c_{rr+1}x_{r+1} + \cdots + c_{rn}x_n) \end{cases} r$$

$$E_1 \begin{cases} 0 = 0 \\ \cdots\cdots \\ \cdots\cdots \\ 0 = 0 \end{cases} m - r$$

Si es $r = n$, el sistema tiene solución única; si es $r < n$, el sistema tiene infinitas soluciones. En este segundo caso, al tomar para las incógnitas x_{r+1}, ..., x_n unos valores cualesquiera, nos encontramos con que hay unos únicos valores para x_1, x_2, ..., x_r que, junto con aquellos, forman una solución del sistema. Para expresar que los valores de las $n - r$ incógnitas x_{r+1}, ..., x_n (incógnitas secundarias) se pueden elegir libremente, se acostumbra a decir, que las soluciones del sistema forman una familia con $n - r$ *grados de libertad*. El valor que toma $n - r$ es fijo, es decir, depende del sistema S pero no del camino que se haya seguido para obtener E_1. Esta afirmación, que quizá le parezca evidente al lector, no se ha demostrado aún; lo haremos, más adelante, cuando volvamos sobre este asunto de los sistemas de ecuaciones.

Al echar una ojeada al sistema E_1, se ve que en él hay r ecuaciones «efectivas» (las r primeras) y que las $m - r$ restantes son identidades, se satisfacen para cualesquiera valores de las incógnitas. Esta situación viene a reflejar lo que ocurre en S: de entre las m ecuaciones de S, va a haber r que son independientes, de manera que cada una añade una nueva exigencia, y las $m - r$ restantes resultarán ser consecuencia de las r primeras. Para expresar esto, diremos que r es el *rango* del sistema S o de la matriz de sus coeficientes. Más adelante veremos que estas consideraciones intuitivas tienen total justificación.

(*) Si no se altera el orden de sus incógnitas, es decir, manteniendo su numeración, el sistema E tendría la siguiente expresión:

$$\begin{cases} x_{i_1} = h_{i_1} - (c_{i_1 r+1}x_{i_{r+1}} + \cdots + c_{i_1 n}x_{i_n}) \\ x_{i_2} = h_{i_2} - (c_{i_2 r+1}x_{i_{r+1}} + \cdots + c_{i_2 n}x_{i_n}) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_{i_r} = h_{i_r} - (c_{i_r r+1}x_{i_{r+1}} + \cdots + c_{i_r n}x_{i_n}) \\ 0 = 0 \\ \cdots\cdots \\ \cdots\cdots \\ 0 = 0 \end{cases}$$

(i_1, i_2, ..., i_r son las r índices que señalan los lugares en que están situadas las cabeceras de las ecuaciones del sistema E)

1.5. SISTEMAS HOMOGENEOS CON MENOS ECUACIONES QUE INCOGNITAS

Sabemos que los sistemas lineales homogéneos tienen siempre solución; estos sistemas pueden ser o compatibles determinados, cuando sólo tienen la solución nula, o compatibles indeterminados y, entonces, tienen infinitas soluciones, pues si tienen una solución también son soluciones suyas todas las proporcionales a aquélla. Pues bien, vamos a considerar aquí una condición suficiente para que los sistemas homogéneos tengan infinitas soluciones:

[010]

Sea H un sistema homogéneo de m ecuaciones lineales con n incógnitas:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array}\right\} \quad [H]$$

Si H tiene menos ecuaciones que incógnitas, o sea, si es $m < n$, entonces H tiene alguna solución no nula y, por tanto, tiene infinitas soluciones(*).

(*) Se supone que los escalares son los elementos de un cuerpo infinito, como $\mathbb{R}$ o como $\mathbb{C}$.

DEMOSTRACION

Aplicando el método de Gauss al sistema H, siempre es posible obtener un sistema escalonado E equivalente al H; el sistema E ha de ser también homogéneo, como resulta obvio. De acuerdo con la discusión de los sistemas escalonados que se hizo en [007] nos encontramos con que, en nuestro caso, como todos los términos independientes son nulos y es $r \leqslant m < n$ (se llama r, como siempre, al número de ecuaciones de E que tienen sus primeros miembros no nulos), resulta que (según se dijo en [007]) E es compatible indeterminado, luego también lo es H.

Como H ha resultado tener más de una solución, H tiene entonces alguna solución no nula; sea ésta la $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Consecuentemente, también son soluciones de H todas las infinitas $(\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n)$ para cualquier escalar λ, como se comprueba trivialmente.

CAPITULO

2

Rango (de vectores y de matrices)

Ya hemos tenido, anteriormente, un encuentro con el rango, cuando hablábamos del proceso de resolución de un sistema de ecuaciones lineales. Recuérdese que allí, al estudiar un sistema de m ecuaciones, desechábamos todas las que eran combinaciones lineales de otras y, a la postre, terminábamos quedándonos, solamente, con aquellas que, en número mínimo, eran independientes unas de otras y tales que de ellas dependían todas las rechazadas. Este número mínimo de ecuaciones era el rango.

Vamos ahora a precisar este concepto y a profundizar en él. Esta tarea la haremos desde los puntos de vista vectorial y matricial.

α VECTORES DE n COMPONENTES

Mas adelante, cuando sea preciso, hablaremos de los espacios vectoriales abstractos con toda la generalidad que nos sea necesaria. Por el momento no nos hace falta tanto; sólo hablaremos aquí de los «vectores de n componentes», situados ya en el terreno de lo abstracto, pero no muy lejos de los vectores elementales que, de seguro, hemos venido utilizando en nuestros estudios anteriores. Hablar aquí de los «vectores de n componentes» nos ha de ayudar, en su momento, a llegar sosegadamente al concepto de espacio vectorial.

A poca Física que hayamos estudiado, habremos visto que esta ciencia está «plagada» de vectores: las fuerzas, las velocidades, los campos eléctricos, etc. La cosa no fue siempre así; los vectores entraron en la Física con no poca resistencia por parte de muchos, incluso de gentes eminentes. A este respecto, bueno es recordar que el gran físico Lord Kelvin llegó a decir que «los vectores nunca han tenido la menor utilidad para nadie».

Todos conocemos ya los vectores geométricos (segmentos orientados del espacio) y sabemos que, recurriendo a las coordenadas, se pueden identificar con las ternas de números reales, o sea, con los vectores de tres componentes reales. Los vectores de n componentes están ya haciéndonos falta, hay que hablar de ellos sin demora. Nótese, a este respecto, que en los apartados anteriores ya se ha recurrido tímidamente a ellos cuando se echó mano de objetos como el $\bar{x} = (x_1, x_2, ..., x_n)$, que nos sirvió para representar a todos y cada uno de los elementos de la sucesión $x_1, x_2, ..., x_n$.

Abundando en la conveniencia de estudiar aquí los vectores de n componentes y no otros más generales, que se aplazan para más adelante, hay que argüir otra razón de importancia: Las propiedades de los vectores de n componentes, que nos va a costar muy poco comprobar, se tomarán como punto de partida para definir vectores en general. Conocer dichas propiedades nos va a permitir entender las razones por las que los vectores abstractos se definen del modo que se hace y no de otra manera.

2.1. EL ESPACIO VECTORIAL K^n

Las siguientes definiciones generalizan, de manera natural y obvia, a las nociones de vectores de 2 y de 3 componentes (que nos son conocidas) así como las de su suma y de producto por un escalar. Con todo ello se obtiene una estructura algebraica que, según se verá más adelante, no es otra que la de espacio vectorial de dimensión n.

DEFINICION

[011]

Dados un número natural n (fijo) y un cuerpo K (en particular $K = \mathbb{R}$ o $K = \mathbb{C}$), a cuyos elementos llamaremos escalares, se llaman «vectores de n componentes» del cuerpo K (en particular, reales o complejas) a los elementos que se especifican en I entre las que hay establecidas las dos operaciones que se detallan en II:

I. Se llaman vectores de n componentes del cuerpo K a las sucesiones o sistemas ordenados de n escalares de K, es decir, a los elementos de K^n. Estos vectores son, pues, los objetos:

$$\bar{u} = (u_1, u_2, ..., u_n)$$

$$\text{donde} \begin{cases} u_1, u_2, ..., u_n \text{ son escalares (del cuerpo } K) \\ u_i \text{ se llama componente } i\text{-ésima de } \bar{u} \end{cases}$$

Los vectores $\bar{u} = (u_1, u_2, ..., u_n)$ y $\bar{v} = (v_1, v_2, ..., v_n)$ son iguales si tienen iguales sus respectivas componentes; esto es, si $u_i = v_i$ para $i = 1, 2, ..., n$.

II. Se llama suma $\bar{u} + \bar{v}$, de dos vectores $\bar{u}$ y $\bar{v}$, y se llama producto $\lambda\bar{u}$, donde λ es un escalar, a los siguientes vectores:

$$\left.\begin{array}{l} \bar{u} = (u_1, u_2, ..., u_n) \\ \bar{v} = (v_1, v_2, ..., v_n) \end{array}\right\} \rightarrow \bar{u} + \bar{v} = (u_1 + v_1, u_2 + v_2, ..., u_n + v_n)$$

$$\left.\begin{array}{l} \bar{u} = (u_1, u_2, ..., u_n) \\ \lambda \text{ es un escalar} \end{array}\right\} \rightarrow \lambda\bar{u} = (\lambda u_1, \lambda u_2, ..., \lambda u_n)$$

EJEMPLO

Así, para los siguientes vectores de 5 componentes:

$$\bar{u} = (3, -2, 0, 2, 4) \quad ; \quad \bar{v} = (2, 1, -1, 2, 3) \quad ; \quad \bar{w} = (0, 7, -3, 2, 3)$$

los vectores $\bar{u}+\bar{v}$, $\bar{u}-\bar{v}$, $-2\bar{w}$ y $2\bar{u}-3\bar{v}+\bar{w}$ son los:

$$\bar{u}+\bar{v}=(5,\ -1,\ -1,\ 4,\ 7) \quad ; \quad -2\bar{w}=(0,\ -14,\ 6,\ -4,\ -6)$$
$$\bar{u}-\bar{v}=(1,\ -3,\ 1,\ 0,\ 1) \quad ; \quad 2\bar{u}-3\bar{v}+\bar{w}=(0,\ 0,\ 0,\ 0,\ 0)$$

OBSERVACIONES

Supongamos que, en una empresa, las citas se comunican dando un vector (a, b, c), donde a es la hora, b es el día y c es el mes. Así (12, 24, 04) es una cita para las 12 del día 24 de abril. Si alguien se equivoca en el orden y comunica que la cita es (04, 24, 12), el citado deberá acudir a las 4 de la madrugada del día de nochebuena.

1. Un vector de n componentes no es un conjunto de n escalares, sino un elemento de K^n. Si las n componentes de un vector se colocan en distinto orden, ya no encarnan al mismo vector; éste pasa a ser otro.
2. Es costumbre denotar a los vectores de distinto modo que a los escalares; para distinguir a aquéllos se suele poner una tilde o una flechita sobre la letra que los simboliza o, también, escribir ésta con letra negrita. Ello, que suele resultar útil, sobre todo al principio, no es algo necesario; es más, cuando se tiene bastante hábito en el manejo de vectores y escalares, puede llegar a ser incómodo y tedioso verse en la obligación de escribir las referidas tildes.
3. Se suele decir que el vector $\lambda\bar{u}$ es *proporcional* al vector $\bar{u}$ y que λ es la constante de proporcionalidad.
4. Si α es un escalar no nulo, «dividir el vector $\bar{u}$ por α» es multiplicar el vector $\bar{u}$ por el escalar $1/\alpha$.
5. *Vectores fila* y *vectores columna*. Para señalar que las componentes de un vector $\bar{u}$ se escriben una detrás de otra, en el mismo renglón, se suele decir que $\bar{u}$ aparece como vector fila. Si las componentes de $\bar{u}$ se escriben en vertical, poniendo cada una debajo de su anterior, se dice que $\bar{u}$ se presenta en forma de vector columna:

$$\bar{u}=\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{o} \quad \bar{u}=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

Con esta notación en forma de columna, las operaciones entre vectores se presentarán en la siguiente forma:

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}+\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}=\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_n+v_n \end{bmatrix} \quad ; \quad \lambda\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}=\begin{bmatrix} \lambda u_1 \\ \lambda u_2 \\ \vdots \\ \lambda u_n \end{bmatrix}$$

PROPIEDADES DE LAS OPERACIONES CON VECTORES

Para los vectores de n componentes y en relación con la suma y el producto por un escalar, se verifica:

[012]

Si $\bar{u}$, $\bar{v}$ y $\bar{w}$ son vectores de n componentes (n escalares de un cuerpo K; en particular, reales o complejos) y si λ y μ son dos escalares, entonces:

1. $(\bar{u} + \bar{v}) + \bar{w} = \bar{u} + (\bar{v} + \bar{w})$
2. $\bar{u} + \bar{v} = \bar{v} + \bar{u}$
3. $\bar{u} + \bar{o} = \bar{u}$, donde $\bar{o} = (0, 0, ..., 0)$ se llama *vector nulo*
4. Para cada $\bar{u}$, existe un vector $-\bar{u}$ tal que $\bar{u} + (-\bar{u}) = \bar{o}$;
   Si $\bar{u} = (u_1, u_2, ..., u_n)$, entonces $-\bar{u} = (-u_1, -u_2, ..., -u_n)$
   Al vector $-\bar{u}$ se le llama opuesto del $\bar{u}$.
5. $\lambda(\bar{u} + \bar{v}) = \lambda\bar{u} + \lambda\bar{v}$
6. $(\lambda + \mu)\bar{u} = \lambda\bar{u} + \mu\bar{u}$
7. $\lambda(\mu\bar{u}) = (\lambda\mu)\bar{u}$
8. $1\bar{u} = \bar{u}$

---

a) $0\bar{u} = \bar{o}$ y $\lambda\bar{o} = \bar{o}$
b) $\lambda\bar{u} = \bar{o} \Rightarrow [\lambda = 0 \text{ o } \bar{u} = \bar{o}]$
c) $(-\lambda)\bar{u} = \lambda(-\bar{u}) = -(\lambda\bar{u})$

Obsérvese que las últimas propiedades se han separado de las primeras, con una raya horizontal; con ello se quiere indicar que, cuando se definan los espacios vectoriales abstractos, las ocho primeras propiedades serán las que caractericen a los vectores y las tres últimas van a ser consecuencias de aquéllas.

COMPROBACION

Las anteriores propiedades se comprueban trivialmente; hagámoslo, como ejercicio, para dos de ellas (la 5 y la *c*):

- Llamando $\bar{u} = (u_1, u_2, ..., u_n)$ y $\bar{v} = (v_1, v_2, ..., v_n)$, para la propiedad 5 se tiene:

$$\begin{aligned}\lambda(\bar{u} + \bar{v}) &= \lambda[(u_1, u_2, ..., u_n) + (v_1, v_2, ..., v_n)] = \\ &= \lambda(u_1 + v_1, u_2 + v_2, ..., u_n + v_n) = \\ &= [\lambda(u_1 + v_1), \lambda(u_2 + v_2), ..., \lambda(u_n + v_n)] = \\ &= (\lambda u_1 + \lambda v_1, \lambda u_2 + \lambda u_2, ..., \lambda u_n + \lambda v_n) = \\ &= (\lambda u_1, \lambda u_2, ..., \lambda u_n) + (\lambda v_1, \lambda v_2, ..., \lambda v_n) = \lambda\bar{u} + \lambda\bar{v}\end{aligned}$$

- La propiedad *b*), en lugar de comprobarla directamente, que no sería difícil, vamos a obtenerla como consecuencia de las primeras propiedades. Nótese que esta propiedad se puede enunciar diciendo que: de ser $\lambda\bar{u} = \bar{o}$ y $\lambda \neq 0$, entonces se verifica que $\bar{u} = \bar{o}$. Así ocurre, ya que por ser $\lambda \neq 0$ existe λ^{-1} y, por tanto:

$$\lambda\bar{u} = \bar{o} \Rightarrow \lambda^{-1}(\lambda\bar{u}) = \lambda\bar{o} \Rightarrow (\lambda^{-1}\lambda)\bar{u} = \bar{o} \Rightarrow 1\bar{u} = \bar{o} \Rightarrow \bar{u} = \bar{o}$$

2.2. DEPENDENCIA E INDEPENDENCIA LINEAL

Vamos ahora a concretar una idea que ya hemos manejado anteriormente y que, en lo que sigue, va a seguir siendo enormemente fructífera; n concepto de «vector que depende linealmente de otros»

COMBINACIONES LINEALES; DEPENDENCIA LINEAL

[013]

Se dice que un vector $\bar{v}$ es una *combinación lineal* de los vectores $\bar{u}_1$, $\bar{u}_2$, ..., $\bar{u}_k$ o que $\bar{v}$ *depende linealmente* de ellos si, para algunos escalares x_1, x_2, ..., x_k, se verifica:

$$\bar{v} = x_1\bar{u}_1 + x_2\bar{u}_2 + \cdots + x_k\bar{u}_k$$

La dependencia lineal es transitiva: si $\bar{v}$ depende linealmente de unos *vectores y cada uno de estos depende linealmente de otros*, entonces $\bar{v}$ depende linealmente de los últimos.

Se dice que un sistema (o conjunto) de vectores $S = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_p\}$ es un sistema *ligado* o *linealmente dependiente* si se verifica una de las siguientes condiciones (I) o (II), que son equivalentes entre sí:

(I) Alguno de los vectores de S depende linealmente de los demás.
(II) Para algunos escalares λ_1, λ_2, ..., λ_p que no son todos nulos es:

$$\lambda_1\bar{v}_1 + \lambda_2\bar{v}_2 + \cdots + \lambda_p\bar{v}_p = \bar{o}$$

DEMOSTRACION

1. Comprobemos en primer lugar la transitividad de la dependencia lineal. Se supone que $\bar{v}$ depende linealmente de los $\bar{u}_1$, $\bar{u}_2$, ..., $\bar{u}_k$ y que cada uno de estos $\bar{u}_i$ depende linealmente de los $\bar{w}_1$, $\bar{w}_2$, ..., $\bar{w}_k$; esto es, para ciertos escalares λ_i y μ_{ij}, se verifica que:

$$\bar{v} = \sum_{i=1}^{h} \lambda_i\bar{u}_i \quad \text{y} \quad \bar{u}_i = \sum_{j=1}^{k} \mu_{ij}\bar{w}_j$$

De aquí se infiere que:

$$\bar{v} = \sum_{i=1}^{h} \lambda_i\left(\sum_{j=1}^{k} \mu_{ij}\bar{w}_j\right) = \sum_{i=1}^{h}\sum_{j=1}^{k} \lambda_i\mu_{ij}\bar{w}_j = \sum_{j=1}^{k}\left(\sum_{i=1}^{h} \lambda_i\mu_{ij}\right)\bar{w}_j = \sum_{j=1}^{k} \alpha_j\bar{w}_j$$

donde se ha llamado

$$\alpha_j = \sum_{i=1}^{h} \lambda_i \mu_{ij}$$

Resulta, pues, que $\bar{v}$ depende linealmente de los $\bar{w}_j$, según había que comprobar.

2. También debemos comprobar que las condiciones I y II son equivalentes; esto es, que I $\Rightarrow$ II y que II $\Rightarrow$ I. Así ocurre, en efecto:

I $\Rightarrow$ II La hipótesis es, ahora, que alguno de los $\bar{v}_i$ depende linealmente de los demás; vamos a suponer que dicho vector es el $\bar{v}_1$ (de no ser así, cambiamos la numeración para que lo sea), es decir, que para algunos escalares λ_2, ..., λ_h es:

$$\bar{v}_1 = \lambda_2 \bar{v}_2 + \lambda_3 \bar{v}_3 + \cdots + \lambda_h \bar{v}_h$$

Pasando $\bar{v}_1$ al segundo miembro y llamando $\lambda_1 = -1$, resulta que

$$\bar{o} = \lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2 + \cdots + \lambda_h \bar{v}_h$$

luego se verifica la condición II, como había que probar, ya que los λ_1, λ_2, ..., λ_h no son todos nulos por ser $\lambda_1 = -1 \neq 0$.

II $\Rightarrow$ I La hipótesis es ahora que $\lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2 + \cdots + \lambda_h \bar{v}_h = \bar{o}$ para algunos escalares λ_i no todos nulos. Podemos, pues, suponer que es $\lambda_1 \neq 0$ (de no ser así, cambiando la numeración se consigue que lo sea), lo que nos permite poner:

$$\bar{v}_1 = \left(-\frac{\lambda_2}{\lambda_1}\right)\bar{v}_2 + \cdots + \left(-\frac{\lambda_h}{\lambda_1}\right)\bar{v}_h$$

es decir, uno de los vectores de S (el $\bar{v}_1$) depende linealmente de los demás, esto es, se verifica la condición I.

EJEMPLOS

1. El vector nulo, $\bar{o}$, es combinación lineal de cualesquiera vectores $\bar{u}_1$, $\bar{u}_2$, ..., $\bar{u}_h$, ya que

$$\bar{o} = 0\bar{u}_1 + 0\bar{u}_2 + \cdots + 0\bar{u}_h$$

Por tanto si $\bar{o}$ es uno de los vectores de un sistema S, entonces S es linealmente dependiente.

2. Sean $\bar{u}_1$, $\bar{u}_2$, $\bar{u}_3$ y $\bar{v}$ los siguientes vectores de 4 componentes:

$$\bar{u}_1 = (2, -1, 5, 1)$$
$$\bar{u}_2 = (-1, 3, -2, 0) \qquad \bar{v} = (1, -8, 1, -1)$$
$$\bar{u}_3 = (3, 1, 8, 2)$$

El vector $\bar{v}$ depende linealmente de los vectores $\bar{u}_1$, $\bar{u}_2$ y $\bar{u}_3$, ya que:

$$\bar{v} = 1\bar{u}_1 - 2\bar{u}_2 - 1\bar{u}_3$$

Nótese que, en este caso, $\bar{v}$ también puede expresarse de otros modos como combinación lineal de los $\bar{u}_1$, $\bar{u}_2$ y $\bar{u}_3$, ya que:

$$\bar{v} = 3\bar{u}_1 - \bar{u}_2 - 2\bar{u}_3 \quad \text{y} \quad \bar{v} = -\bar{u}_1 - 3\bar{u}_2 + 0\bar{u}_3$$

Señalemos, de paso, que la última igualdad nos permite afirmar que $\bar{v}$ es una combinación lineal de los $\bar{u}_1$ y $\bar{u}_2$ (solamente).

3. Sea $S = \{\bar{a}, \bar{b}, \bar{c}\}$ el sistema formado por los siguientes vectores

$$\bar{a} = (1, 4, -2) \quad , \quad \bar{b} = (4, 1, 7) \quad \text{y} \quad \bar{c} = (2, -1, 5)$$

Este sistema S es un sistema ligado ya que $2\bar{a} - 3\bar{b} + 5\bar{c} = \bar{o}$. Nótese que, para cualesquiera que sean los vectores $\bar{d}$ y $\bar{e}$ (de tres componentes), el sistema $S' = \{\bar{a}, \bar{b}, \bar{c}, \bar{d}, \bar{e}\}$ también es ligado, ya que

$$2\bar{a} - 3\bar{b} + 5\bar{c} + 0\bar{d} + 0\bar{e} = \bar{o}$$

PROBLEMA

Pruébese que, si $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_p$ son p vectores de n componentes y si es $p > n$, entonces $S = \{\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_p\}$ es un sistema ligado.

RESOLUCION

Denotemos a los vectores dados, $\bar{a}_j$, del siguiente modo:

$$\bar{a}_j = (a_{1j}, a_{2j}, \ldots, a_{nj}) \quad \text{para} \quad j = 1, 2, \ldots, p$$

Hay que comprobar que existen unos ciertos escalares $x_1, x_2, \ldots, x_p$ no todos nulos y tales que

$$x_1\bar{a}_1 + x_2\bar{a}_2 + \cdots + x_p\bar{a}_p = \bar{o}$$

es decir, de manera que:

$$x_1(a_{11}, a_{21}, \ldots, a_{n1}) + x_2(a_{12}, a_{22}, \ldots, a_{n2}) + \cdots +$$
$$+ x_p(a_{1p}, a_{2p}, \ldots, a_{np}) = (0, 0, \ldots, 0)$$

Esta igualdad vectorial equivale a las n ecuaciones (escalares) que resultan de

igualar a cero cada una de las componentes de la suma del primer miembro, esto es, al sistema:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{np}x_p = 0 \end{array}\right\} \quad [H]$$

Como H es un sistema homogéneo que tiene mayor número de incógnitas (p) que de ecuaciones (n), según ya sabemos (véase [010]) dicho sistema tiene alguna solución no nula, que es lo que había que comprobar.

INDEPENDENCIA LINEAL

Si unos vectores no forman un sistema ligado, diremos que son linealmente independientes; más precisamente:

[014]

Se dice que un sistema (o conjunto) de vectores $S = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_p\}$ es un sistema *libre* o *linealmente independiente* si se verifica una de las siguientes condiciones (A) o (B), que son equivalentes entre sí:

(A) El sistema S no es linealmente dependiente, es decir, ninguno de los vectores de S depende linealmente de los demás.

(B) La única combinación lineal de vectores de S que es nula es la que tiene todos sus coeficientes nulos; esto es:

$$\lambda_1\bar{v}_1 + \lambda_2\bar{v}_2 + \cdots + \lambda_p\bar{v}_p = \bar{o} \;\Rightarrow\; \lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$$

($\lambda_1, \lambda_2, ..., \lambda_p$ escalares).

COMPROBACION

Si comparamos las condiciones (A) y (B) con las (I) y (II) de la dependencia lineal (véase [013]), nos encontramos con que (A) es la negación de (I) y (B) es la negación de (II). Como (I) y (II) son equivalentes, según ya se demostró, concluimos de ello que también son equivalentes sus negaciones (A) y (B).

EJEMPLO

Los vectores $\bar{a}$, $\bar{b}$ y $\bar{c}$ (de 5 componentes):

$$\bar{a} = (3, 1, 7, -5, 4) \quad , \quad \bar{b} = (0, 2, 4, -3, -2) \quad \text{y} \quad \bar{c} = (0, 0, -6, 5, 1)$$

forman un sistema libre, ya que la relación $\alpha\bar{a} + \beta\bar{b} + \gamma\bar{c} = \bar{o}$ sólo se verifica para $\alpha = \beta = \gamma = 0$. Así ocurre, en efecto, ya que la citada relación equivale al sistema:

$$\left.\begin{aligned} 3\alpha = 0 \\ \alpha + 2\beta = 0 \\ 7\alpha + 4\beta - 6\gamma = 0 \\ -5\alpha - 3\beta + 5\gamma = 0 \\ 4\alpha - 2\beta + \gamma = 0 \end{aligned}\right\}$$ que sólo tiene (obviamente) la solución $\alpha = 0,\ \beta = 0,\ \gamma = 0.$

β RANGO DE UN SISTEMA DE VECTORES

Dado un conjunto de vectores no nulos, estamos interesados en estudiar si hay algunos de ellos tales que satisfagan a los dos requerimientos siguientes: 1) que formen un sistema independiente; y 2) que los demás vectores dados dependan linealmente de ellos. Vamos a probar que tales sistemas existen y que todos ellos tienen el mismo número de vectores, que se llama rango del conjunto de vectores dado.

2.3. RANGO. OPERACIONES ELEMENTALES

TEOREMA (EXISTENCIA DEL RANGO)

[015]

Para cualquiera que sea el sistema $S = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_p\}$ de p vectores, no todos nulos, se verifica que:

(I) Hay algún sistema S_0 de vectores de S tal que: 1) S_0 es linealmente independiente; y 2) todos los demás vectores de S dependen linealmente de los de S_0.

(II) Todos los sistemas S_0, de vectores de S, que satisfacen a los requerimientos 1) y 2) tienen el mismo número de vectores.

Para cada sistema $S = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_p\}$ de vectores no nulos, existe pues un número r tal que: a) hay algún sistema formado por r vectores de S que es linealmente independiente; y b) todo sistema formado por más de r vectores de S es linealmente dependiente. Este número r es, entonces, el mayor número de vectores linealmente independientes que hay en S; a él se le llama *rango* de S y se pone:

$$r = \text{rang } S$$

DEMOSTRACION

Hemos de comprobar la veracidad de las dos proposiciones (I) y (II). Las demás afirmaciones del enunciado son consecuencias obvias de (I) y (II).

(I) Esta proposición la comprobaremos por inducción sobre p; esto es, verificaremos que es cierta para $p = 1$ y probaremos que, de ser cierta para $p = m$, entonces también lo es para $p = m + 1$. Esta proposición (I) es evidentemente cierta para $p = 1$ ya que, como ahora es $S = \{\bar{v}_1\}$ con $\bar{v}_1 \neq \bar{o}$, la proposición se satisface obviamente para $S_0 = S$. Suponiendo ahora que la proposición (I) es cierta para los sistemas S que tengan $p = m$ vectores, debemos comprobar que también lo es si S tiene $p = m + 1$ vectores. Consideremos, para ello, los dos casos que pueden presentarse (los $m + 1$ vectores son independientes o son dependientes). Si los $m + 1$ vectores de S son independientes, entonces la propiedad se verifica para $S_0 = S$. Si los $m + 1$ vectores de S son linealmente dependientes, entonces uno de ellos depende linealmente de los demás; llamemos a aquel $\bar{v}_{m+1}$, con lo que los demás son $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m$. En este caso, como el sistema $S' = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m\}$ tiene m vectores, para él existe el sistema S'_0 que asegura la proposición (I); esta proposición (I) se verifica, entonces, (para S) si se toma $S_0 = S'_0$, ya que este S'_0 es independiente y todos los vectores de S (incluyendo el $\bar{v}_{m+1}$) dependen, obviamente, de los de S'_0.

(II) Procederemos por reducción al absurdo; esto es, partiendo de que la proposición es falsa, encontraremos una contradicción. Supongamos, pues, que hubiera dos sistemas S_0, que llamaremos S'_0 y S''_0, que cumplieran lo exigido en (I) y tuvieran distinto número de vectores. Si denotamos

$$S'_0 = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_r\} \quad \text{y} \quad S''_0 = \{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_s\}$$

lo que se ha admitido es que r y s son distintos; supongamos, por ejemplo, que es $r < s$. Pues bien, vamos a comprobar que por verificarse que es $r < s$ y que para S'_0 se cumple lo exigido en (I), el sistema S''_0 no puede ser linealmente independiente; esta contradicción conduce a la veracidad de (II). Vamos, pues, a comprobar que existen algunos escalares $x_1, x_2, \ldots, x_s$ no todos nulos y tales que

$$x_1\bar{w}_1 + x_2\bar{w}_2 + \cdots + x_s\bar{w}_s = \bar{o} \quad \text{,o sea} \quad \sum_{i=1}^{s} x_i\bar{w}_i = \bar{o} \qquad [1]$$

Como todos los vectores de S, y en particular los de S''_0, dependen linealmente de los vectores de S'_0, resulta que para unos ciertos escalares a_{ji} se verifica que

$$\bar{w}_i = \sum_{j=1}^{r} a_{ji}\bar{v}_j \quad (\text{para } i = 1, 2, \ldots, s)$$

Recurriendo a estas igualdades, la relación [1] puede expresarse del siguiente modo:

$$\bar{o} = \sum_{i=1}^{s} x_i\bar{w}_i = \sum_{i=1}^{s} x_i\left(\sum_{j=1}^{r} a_{ji}\bar{v}_j\right) = \sum_{i=1}^{s}\sum_{j=1}^{r} x_i a_{ji}\bar{v}_j = \sum_{j=1}^{r}\sum_{i=1}^{s} x_i a_{ji}\bar{v}_j = \sum_{j=1}^{r}\left(\sum_{i=1}^{s} a_{ji}x_i\right)\bar{v}_j$$

esto es:

$$\left(\sum_{i=1}^{s} a_{1i}x_i\right)\bar{v}_1 + \left(\sum_{i=1}^{s} a_{2i}x_i\right)\bar{v}_2 + \cdots \left(\sum_{i=1}^{s} a_{ri}x_i\right)\bar{v}_r = \bar{o} \qquad [2]$$

Como $\{\bar{v}_1, \bar{v}_2, ..., \bar{v}_r\}$ es un sistema linealmente independiente, la relación [2] se verifica si y sólo si son nulos todos los escalares que figuran en ella; esto es, si:

$$\sum_{i=1}^{s} a_{ji}x_i = 0 \quad \text{para} \quad j = 1, 2, ..., r \qquad [3]$$

La relación [3] es un sistema lineal homogéneo de r ecuaciones con las s incógnitas $x_1, x_2, ..., x_s$. Como se ha supuesto que es $r < s$, resulta que el sistema [3] tiene alguna solución no nula (según se probó en [010]). Como [3] es equivalente a [1], nos encontramos con que [1] se verifica para algunos escalares $x_1, x_2, ..., x_s$ no todos nulos, como queríamos comprobar.

OBSERVACION

Para hallar el rango de un sistema de vectores habrá que ir desechando aquellos vectores que resulten ser combinación lineal de los que van quedando, hasta que éstos formen un sistema independiente; el número de vectores de éste es el rango buscado. Conviene proceder, para ello, de una manera metódica, como por ejemplo:

Sea S un sistema, de p vectores no todos nulos, del que se quiere hallar el rango. Tomemos un vector no nulo de S; sea éste el $\bar{v}_1$. Si todos los vectores de S dependen linealmente de $\bar{v}_1$, entonces el rango es 1. Si en S hay algún vector que no dependa linealmente de $\bar{v}_1$, al que llamaremos $\bar{v}_2$, y los demás vectores dependen de los $\bar{v}_1$ y $\bar{v}_2$, entonces el rango de S es 2. Si en S hay un vector, que llamaremos $\bar{v}_3$, que no depende linealmente de los $\bar{v}_1$ y $\bar{v}_2$ y todos los demás vectores dependen linealmente de $\bar{v}_1$, $\bar{v}_2$ y $\bar{v}_3$, entonces el rango de S es 3. Prosiguiendo de este modo, se llega a encontrar (a lo más es p etapas) el rango de S.

EJERCICIO

Sean $S = \{\bar{u}_1, \bar{u}_2, ..., \bar{u}_p\}$ y $S' = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_k\}$ dos sistemas de vectores. Sabiendo que S es un sistema independiente y que los vectores de S' dependen linealmente de los de S, comprobar que rang $S' \leq p$.

RESOLUCION

El sistema $S \cup S' = \{\bar{u}_1, ..., \bar{u}_p, \bar{v}_1, ..., \bar{v}_n\}$ tiene rango p, ya que sus p primeros vectores son linealmente independientes y todos los restantes dependen linealmente de ellos. El rango de S' no puede ser mayor que p, pues si lo fuese, habría más de p vectores de S', y por tanto, de $S \cup S'$, que serían linealmente independientes, lo que es falso ya que el rango de $S \cup S'$ es p.

EJEMPLO

Sea $S = \{\bar{a}, \bar{b}, \bar{c}, \bar{d}\}$ el sistema formado por los vectores:

$$\bar{a} = (1, 0, 0, 0, 1)$$
$$\bar{b} = (0, 2, 0, 2, 0)$$
$$\bar{c} = (3, 0, 3, 0, 3)$$
$$\bar{d} = (1, 2, 3, 2, 1)$$

Los vectores $\bar{a}$, $\bar{b}$ y $\bar{c}$ forman un sistema independiente, ya que la relación $\alpha\bar{a} + \beta\bar{b} + \gamma\bar{c} = \bar{o}$ equivale a:

$$\alpha + 3\gamma = 0 \quad ; \quad 2\beta = 0 \quad ; \quad 3\gamma = 0$$

cuya única solución es la $\alpha = \beta = \gamma = 0$. Por otra parte, $\bar{d}$ es una combinación lineal de las $\bar{a}$, $\bar{b}$ y $\bar{c}$, ya que $\bar{d} = -2\bar{a} + \bar{b} + \bar{c}$. Por tanto, el rango de S es rang $S = 3$.

RANGO Y OPERACIONES ELEMENTALES

Igual que ocurría con los sistemas de ecuaciones lineales, también aquí las operaciones elementales juegan un papel importante. Allí (véase [004]), al realizar operaciones elementales en un sistema de ecuaciones lineales, no variaban las soluciones de éste. Aquí vamos a comprobar que, al realizar operaciones elementales en un sistema de vectores, no varía su rango.

[016]

Propiedad fundamental del rango: el rango de un sistema de vectores no se altera si se realizan en él cualesquiera *operaciones elementales*. Se llaman operaciones elementales a las siguientes manipulaciones:

1. Intercambiar el orden con el que figuran los vectores en el sistema.
2. Multiplicar uno de los vectores por cualquier escalar no nulo.
3. Sumarle, a uno de los vectores, otro cualquiera de ellos.
4. Aplicar, reiteradamente, cualesquiera de las operaciones anteriores (en particular, sumarle, a un vector, cualquier combinación lineal de los demás).

NOTA: Hay quienes consideran también como operación elemental (cosa que no haremos nosotros) a la siguiente: suprimir un vector que sea combinación lineal de otros vectores.

COMPROBACION

La conservación del rango al realizar operaciones elementales del tipo 1 es evidente; para la operación 4, como es consecuencia de las tres anteriores, no hay nada que demostrar. Sólo habrá que comprobar la conservación del rango frente a las operaciones 2 y 3. Estas dos propiedades podemos englobarlas en

una sola que diga: el rango de un sistema $S = \{\bar{u}, \bar{v}, \bar{w}, ...\}$ es igual al del sistema $S' = \{\bar{u}', \bar{v}, \bar{w}, ...\}$, donde $\bar{u}' = \alpha\bar{u} + \beta\bar{v}$ para $\alpha \neq 0$ (si es $\beta = 0$, se obtiene la operación 2; si $\alpha = \beta = 1$, se obtiene la operación 3). Para demostrar esto, recurriremos al rango r del sistema $S_0 = \{\bar{v}, \bar{w}, ...\}$ y distinguiremos dos casos:

a) El vector $\bar{u}$ depende linealmente de S_0. En este supuesto, es rang $S = r$. Como es evidente que, ahora, $\bar{u}' = \alpha\bar{u} + \beta\bar{v}$ también depende linealmente de S_0, resulta que rang $S' = r$. Por tanto, S y S' tienen igual rango.

b) El vector $\bar{u}$ no depende linealmente de S_0. En este supuesto, es rang $S = r + 1$. Ahora, $\bar{u}'$ tampoco depende linealmente de S_0 pues si así no fuese, como es $\bar{u} = (1/\alpha)\bar{u}' + (\beta/\alpha)\bar{v}$, resultaría que $\bar{u}$ dependería linealmente de S_0, cuando se ha supuesto lo contrario. Resulta de ello que S' tiene rango $r + 1$, luego S y S' tienen, también en este caso, igual rango.

2.4. CALCULO DEL RANGO DE UN SISTEMA DE VECTORES

Para hallar el rango de un sistema S, es aconsejable proceder sistemáticamente, aplicando adecuadas operaciones elementales, hasta obtener un sistema en el que hallar el rango sea algo elemental. El rango de S, por ser igual al de este último, será ya conocido.

PROPOSICION

[017]

Dado $S = \{\bar{a}, \bar{b}, ..., \bar{e}\}$, sistema formado por los vectores de n componentes:

$$\left.\begin{array}{l} \bar{a} = (a_1, a_2, ..., a_n) \\ \bar{b} = (b_1, b_2, ..., b_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \bar{c} = (c_1, c_2, ..., c_n) \end{array}\right\} \quad \text{sea} \quad M = \begin{bmatrix} a_1 & a_2 & ... & a_n \\ b_1 & b_2 & ... & b_n \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ c_1 & c_2 & ... & c_n \end{bmatrix}$$

Realizar operaciones elementales en S equivale a hacer las operaciones en las filas de la matriz M. Utilizando como pivotes (véase [009]) a ciertos elementos de M, elegidos de modo idóneo para realizar adecuadas operaciones elementales en las filas de M, siempre es posible transformar S en un sistema E que sea escalonado (véase [006]), esto es, tal que cada vector (a partir del 2.º) comience con una sucesión de ceros que tiene algún cero más (uno como mínimo) que la fila anterior. El rango de E, que es igual al de S, es el número de vectores no nulos que contiene.

COMPROBACION

Respecto de como conseguir el sistema escalonado E del que se habla en el enunciado, no hay más que seguir el camino, ya conocido, para transformar M en una matriz escalonada aplicando operaciones elementales a sus filas (véase, en [009], lo dicho al describir el método de Gauss).

Sólo tenemos, pues, que comprobar que el rango de E es igual al número r de vectores no nulos que contiene. Comencemos recordando que E es de la forma

$$E = \{\bar{u}_1, \bar{u}_2, ..., \bar{u}_r, \bar{o}, ..., \bar{o}\}$$

donde la primera componente no nula de cada $\bar{u}_i$, a la que denotaremos por u_{i1}, ocupa algún lugar más a la derecha que la de $\bar{u}_{i-1}$ (*). Nótese que lo que debemos comprobar es que $\bar{u}_1, \bar{u}_2, ..., \bar{u}_r$ son linealmente independientes. Así ocurre, en efecto, ya que la relación $\alpha_1\bar{u}_1 + \alpha_2\bar{u}_2 + \cdots + \alpha_r\bar{u}_r = \bar{o}$ (donde los α_i son escalares) equivale al sistema

$$\left.\begin{aligned} &\alpha_1 u_{11} = 0 \\ &\alpha_1 * + \alpha_2 u_{21} = 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\alpha_1 * + \alpha_2 * + \cdots + \alpha_r u_{r1} = 0 \end{aligned}\right\}$$

(con los $*$ se denotan a las restantes componentes de los vectores $\bar{u}_i$) y este sistema sólo tiene, obviamente, la solución nula $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$.

EJEMPLO

Siguiendo el procedimiento antes descrito. calculemos el rango del sistema $S = \{\bar{a}, \bar{b}, \bar{c}, \bar{d}, \bar{e}\}$, siendo:

$$\begin{aligned} \bar{a} &= (1,\ 3,\ 2,\ -1,\ 4) \\ \bar{b} &= (1,\ 4,\ 1,\ \ 1,\ 3) \\ \bar{c} &= (2,\ 5,\ 5,\ \ 1,\ 7) \\ \bar{d} &= (2,\ 4,\ 5,\ -1,\ 8) \\ \bar{e} &= (3,\ 8,\ 7,\ \ 0,\ 11) \end{aligned}$$

Realizando sucesivas operaciones elementales en las filas de la matriz cuyos vectores fila son los $\bar{a}, \bar{b}, \bar{c}, \bar{d}, \bar{e}$, se obtiene:

$$\begin{array}{lcccccc} \bar{a}' = \bar{a} & \rightarrow & 1 & 3 & 2 & -1 & 1 & 4 \\ \bar{b}' = \bar{b} - \bar{a} & \rightarrow & 0 & 1 & -1 & 2 & 1 & -1 \\ \bar{c}' = \bar{c} - 2\bar{a} & \rightarrow & 0 & -1 & 1 & 3 & 1 & -1 \\ \bar{d}' = \bar{d} - 2\bar{a} & \rightarrow & 0 & -2 & 2 & 1 & 0 & 0 \\ \bar{e}' = \bar{e} - 3\bar{a} & \rightarrow & 0 & -1 & 1 & 3 & 1 & -1 \end{array}$$

(*) Los vectores $\bar{u}_i$ son del tipo $\bar{u}_i = (0, ..., 0, u_{i1}, *, ..., *)$; los $*$ son escalares cualesquiera.

$$
\begin{array}{lcrrrrrr}
\bar{a}'' = \bar{a}' & \to & 1 & 3 & 2 & -1 & 1 & 4 \\
\bar{b}'' = \bar{b}' & \to & 0 & -1 & -1 & 2 & 1 & -1 \\
\bar{c}'' = \bar{c}' + \bar{b}' & \to & 0 & 0 & 0 & 5 & 2 & -2 \\
\bar{d}'' = \bar{d}' + 2\bar{b}' & \to & 0 & 0 & 0 & 5 & 2 & -2 \\
\bar{e}'' = \bar{e}' - \bar{c}' & \to & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

$$
\begin{array}{lcrrrrrr}
\bar{a}''' = \bar{a}'' & \to & 1 & 3 & 2 & -1 & 1 & 4 \\
\bar{b}''' = \bar{b}'' & \to & 0 & 1 & -1 & 2 & 1 & -1 \\
\bar{c}''' = \bar{c}'' & \to & 0 & 0 & 0 & 5 & 2 & -2 \\
\bar{d}''' = \bar{d}'' - \bar{c}'' & \to & 0 & 0 & 0 & 0 & 0 & 0 \\
\bar{e}''' = \bar{e}'' & \to & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

El rango de este último sistema es 3, pues $\bar{a}'''$, $\bar{b}'''$ y $\bar{c}'''$ son independientes y $\bar{d}''' = \bar{o}$ y $\bar{e}''' = \bar{o}$; por tanto, el rango del sistema dado, S, es 3.

γ RANGO DE UNA MATRIZ

2.5. EXISTENCIA Y CALCULO DEL RANGO

Una matriz que tenga m filas y n columnas puede concebirse como un sistema (sucesión) de m vectores (vectores fila) de n componentes cada uno de ellos; la matriz también es un sistema (sucesión) de n vectores (vectores columna) de m componentes cada uno. Concretando, para la matriz

$$
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn}
\end{bmatrix}
$$

los a_{ij} son escalares (de un cuerpo K; en particular $a_{ij} \in \mathbb{R}$ o $a_{ij} \in \mathbb{C}$), para $i = 1, 2, ..., m$ y $j = 1, 2, ..., n$. Los vectores fila y los vectores columna son:

$$\text{(vector fila } i\text{-ésima)} \to (a_{i1}, a_{i2}, ..., a_{in})$$

(para cada uno de los valores $i = 1, 2, ..., m$ se obtiene un vector fila de A, que tiene n componentes)

$$\text{(vector columna } j\text{-ésima)} \rightarrow \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

(para cada uno de los valores $j = 1, 2, ..., n$ se obtiene un vector columna, de A, que tiene m componentes).

Pues bien, los vectores fila, de A, tienen el mismo rango que los vectores columna; más exactamente:

TEOREMA DEL RANGO

[018]

Para cualquiera que sea la matriz $A = [a_{ij}]$, de m filas y n columnas (esto es, de tamaño $m \times n$), se verifica que el rango del sistema de sus m vectores fila (que se llama rango de filas de A) es igual al rango del sistema de sus n vectores columna (que se llama rango de columnas de A). A cualquiera de estos dos rangos, iguales entre sí, se le llama *rango de la matriz A*.

DEMOSTRACION

Llamamos r y r' a los rangos de filas y de columnas de la matriz A. Vamos a comprobar que es $r' \leqslant r$ y $r \leqslant r'$, de lo que se infiere que r y r' son iguales. Solamente comprobaremos que es $r' \leqslant r$, pues la relación $r \leqslant r'$ se pruebe de modo totalmente análogo.

Si el rango de filas de A es r, entonces hay r filas de A que son linealmente independientes y las restantes filas dependen de ellas. Cambiando el orden de las filas, si fuera necesario, se puede suponer que las r primeras filas de A son independientes. Llamando $\bar{f}_i$ al vector fila de lugar i, se sabe, pues, que para $i = r+1, r+2, ..., m$, la fila $\bar{f}_i$ es combinación lineal de las $\bar{f}_1, \bar{f}_2, ..., \bar{f}_r$; esto es, para ciertos escalares α_{ih}, se verifica que

$$\bar{f}_i = \alpha_{i1}\bar{f}_1 + \alpha_{i2}\bar{f}_2 + \cdots + \alpha_{ir}\bar{f}_r \quad (\text{para } i = r+1, r+2, ..., m)$$

Esto equivale a poner que las componentes del vector del primer miembro son, respectivamente, iguales a las de la suma del segundo miembro, es decir, a:

$$a_{ij} = \alpha_{i1}a_{1j} + \alpha_{i2}a_{2j} + \cdots + \alpha_{ir}a_{rj} = \sum_{h=1}^{r} \alpha_{ih}a_{hj}$$

(para $i = r+1, r+2, ..., m$ y $j = 1, 2, ..., n$). Por tanto, la columna j-ésima de A, que llamaremos $\bar{c}_j$, se puede poner en la forma:

$$\bar{c}_j = (a_{1j}, a_{2j}, ..., a_{rj}, a_{r+1j}, ..., a_{mj}) =$$

$$= \left(a_{1j}, a_{2j}, ..., a_{rj}, \sum_{h=1}^{r} \alpha_{r+1h} a_{hj}, ..., \sum_{h=1}^{r} \alpha_{mh} a_{hj}\right) =$$

$$= a_{1j}(1, 0, ..., 0, \alpha_{r+1\,1}, ..., \alpha_{m1}) +$$

$$+ a_{2j}(0, 1, ..., 0, \alpha_{r+2\,2}, ..., \alpha_{m2}) +$$

$$.....$$

$$+ a_{rj}(0, 0, ..., 1, \alpha_{r+1\,r}, ..., \alpha_{mr}) =$$

$$= a_{1j}\bar{v}_1 + a_{2j}\bar{v}_2 + \cdots + a_{rj}\bar{v}_r$$

Resulta entonces que todos los vectores columnas de A dependen linealmente de los r vectores que intervienen en esta última suma (que los hemos llamado $\bar{v}_1, \bar{v}_2, ..., \bar{v}_r$) y, por ello, el rango r' de columnas de A no puede ser mayor que el número total r de vectores $\bar{v}_k$, para $k = 1, 2, ..., r$, (véase el ejercicio que sigue a [015], en la página 32); esto es, se verifica que $r' \leqslant r$, como se quería comprobar.

CALCULO DEL RANGO DE UNA MATRIZ

De lo dicho hasta ahora, primero al relacionar el rango de un sistema de vectores con las operaciones elementales (véase [016]), después cuando se sistematizó el cálculo del rango de un sistema de vectores (véase [017]) y, también, al definir los rangos de filas y de columnas de una matriz, que resultaron ser iguales, de todo ello, se desprenden las siguientes conclusiones:

[019]

Para cualquiera que sea la matriz $A = [a_{ij}]$, de m filas y n columnas, se verifica que:

1.° El rango de filas (o de columnas) de A no se altera si se realizan en las filas (o las columnas) de A cualesquiera operaciones elementales.

2.° Utilizando como pivote (véase [009]) a ciertos elementos de A, elegidos de modo idóneo para realizar adecuadas operaciones elementales en las filas (en las columnas) de A, siempre es posible transformar A en una matriz escalonada (escalonada por columnas) E. El rango de filas (de columnas) de la matriz E, que es igual al de A, es el número de filas (de columnas) no nulas que contiene.

EJEMPLO

Como ejemplo, de aplicación inmediata del método descrito, calculemos los rangos de filas y de columnas de la siguiente matriz A, que deberán resultarnos iguales:

$$A = \begin{bmatrix} 1 & 3 & -2 & 1 & 2 \\ 2 & 5 & 1 & -3 & 5 \\ -1 & 3 & 2 & -3 & -3 \\ 3 & -4 & -1 & 2 & 9 \end{bmatrix}$$

Para el rango de filas, al realizar las operaciones elementales, en filas, que se indican a continuación, se llega a:

$$\begin{array}{l} 1.^{a\prime} = 1.^{a} \\ 2.^{a\prime} = 2.^{a} - 2 \cdot 1.^{a} \\ 3.^{a\prime} = 3.^{a} + 1.^{a} \\ 4.^{a\prime} = 4.^{a} - 3 \cdot 1.^{a} \end{array} \rightarrow \begin{bmatrix} 1 & 3 & -2 & 1 & 2 \\ 0 & -1 & 5 & -5 & 1 \\ 0 & 6 & 0 & -2 & -1 \\ 0 & -13 & 5 & -1 & 3 \end{bmatrix} = A'$$

$$\begin{array}{l} 1.^{a\prime\prime} = 1.^{a\prime} \\ 2.^{a\prime\prime} = 2.^{a\prime} \\ 3.^{a\prime\prime} = 3.^{a\prime} + 6 \cdot 2.^{a\prime} \\ 4.^{a\prime\prime} = 4.^{a\prime} - 13 \cdot 2.^{a\prime} \end{array} \rightarrow \begin{bmatrix} 1 & 3 & -2 & 1 & 2 \\ 0 & -1 & 5 & -5 & 1 \\ 0 & 0 & 30 & -32 & 5 \\ 0 & 0 & -60 & 64 & -10 \end{bmatrix} = A''$$

$$\begin{array}{l} 1.^{a\prime\prime\prime} = 1.^{a\prime\prime} \\ 2.^{a\prime\prime\prime} = 2.^{a\prime\prime} \\ 3.^{a\prime\prime\prime} = 3.^{a\prime\prime} \\ 4.^{a\prime\prime\prime} = 4.^{a\prime\prime} + 2 \cdot 3.^{a\prime\prime} \end{array} \rightarrow \begin{bmatrix} 1 & 3 & -2 & 1 & 2 \\ 0 & -1 & 5 & -5 & 1 \\ 0 & 0 & 30 & -32 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = A''' = E$$

Como la matriz escalonada E tiene tres filas no nulas, resulta que el rango de filas de E y, por tanto de A, es 3.

Para el rango de columnas, al realizar las operaciones elementales, en columnas, que se indican a continuación, se obtiene:

$$\begin{array}{l} 1.^{a\prime} = 1.^{a} \\ 2.^{a\prime} = 2.^{a} - 3 \cdot 1.^{a} \\ 3.^{a\prime} = 3.^{a} + 2 \cdot 2.^{a} \\ 4.^{a\prime} = 4.^{a} - 1.^{a} \\ 5.^{a\prime} = 5.^{a} - 2 \cdot 2.^{a} \end{array}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 5 & -5 & 1 \\ -1 & 6 & 0 & -2 & -1 \\ 3 & -13 & 5 & -1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 \\ -1 & 6 & 30 & -32 & 5 \\ 3 & -13 & -60 & -64 & -10 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 \\ -1 & 6 & 1 & 0 & 0 \\ 3 & -13 & -2 & 0 & 0 \end{bmatrix}$$

esta última matriz escalonada tiene tres columnas no nulas, luego su rango de columnas, que es igual al de A, vale 3.

OBSERVACION

Según lo que venimos diciendo, al realizar operaciones elementales se conserva el rango. Ello no significa que el rango debe mantenerse siempre que se sustituyan filas o columnas por otras que sean combinaciones lineales de aquéllas; al hacer esto, el rango no aumenta, pero puede disminuir. Así, por ejemplo, consideremos una matriz A formada por tres columnas independientes, $\bar{c}_1$, $\bar{c}_2$ y $\bar{c}_3$, con lo que rang $A = 3$; sea A' la matriz cuyas columnas son $\bar{c}'_1 = -\bar{c}_1 + \bar{c}_2$, $\bar{c}'_2 = -\bar{c}_1 + \bar{c}_3$ y $\bar{c}'_3 = 2\bar{c}_1 - \bar{c}_2 - \bar{c}_3$. Las tres columnas de A' son linealmente dependientes, pues $\bar{c}'_1 + \bar{c}'_2 + \bar{c}'_3 = \bar{o}$, luego rang $A' < 3$, con lo que rang $A' <$ rang A.

2.6. MATRICES EQUIVALENTES

Más adelante, cuando se habla de las matrices de las aplicaciones lineales, se volverá sobre este concepto de equivalencia de matrices y, entonces, se le dará un enfoque nuevo y de mayor alcance. No obstante, la estrecha relación que existe entre el rango y la equivalencia de matrices, hace aconsejable que sea ahora, y no más adelante, cuando se introduzca este concepto.

[020]

Dos matrices del mismo tamaño $m \times n$ (ambas con m filas y n columnas) se dice que son *matrices equivalentes* si tienen el mismo rango. Para el conjunto de las matrices $m \times n$, la equivalencia es (efectivamente) una relación de equivalencia; es decir, es reflexiva, simétrica y transitiva. Hay tantas clases de esta equivalencia como posibles valores del rango r, los cuales son todos los números naturales comprendidos entre $r = 0$ y $r = \text{mín}\ (m, n)$, ambos incluive.

De entre todas las matrices $m \times n$ que tienen rango r, interesa destacar la que hemos llamado C_r, que tiene todos sus elementos nulos excepto los r primeros elementos de su diagonal (*) que valen 1. Esta matriz C_r se llama representante canónico o *matriz canónica de equivalencia* de las matrices $m \times n$ que tienen rango r.

Aplicando adecuadas operaciones elementales, en filas y en columnas, cualquier matriz $m \times n$ de rango r se puede transformar en C_r.

$$C_r = \begin{bmatrix} 1 & & & O & & \\ & 1 & & & & \\ & & \ddots & & r & O \\ O & & & 1 & & \\ & & & & & \\ r & & & & O & \\ O & & & & & \end{bmatrix}$$

(*) La diagonal de una matriz está formada por los elementos que ocupan los lugares (1, 1), (2, 2), (3,3), ...

COMPROBACIONES

- Es trivial comprobar que la equivalencia de matrices es reflexiva (toda matriz tiene el mismo rango que ella misma), es simétrica (si rang A = rang B, entonces rang B = rang A) y es transitiva (si rang A = rang B, y rang B = rang C, entonces rang A = rang C).
- La matriz C_r tiene rango r pues sus r primeras filas (y columnas) son linealmente independientes y todas las restantes son nulas.
- Según ya sabemos (véase [019]), realizando las oportunas operaciones elementales, a una matriz $m \times n$ que tenga rango r, es posible obtener una matriz escalonada con r filas no nulas, las cuales tienen a 1 como elemento de cabecera, esto es, como primer elemento no nulo. Utilizando como pivotes a estas r cabeceras y aplicando operaciones elementales en columnas, se obtiene una matriz cuyos únicos elementos no nulos son dichas r cabeceras, que valen 1. Sin más que permutar ahora, si ello fuera preciso, algunas columnas, se obtiene ya la matriz C_r.

CAPITULO

3

Operaciones con matrices; matriz inversa

Es ahora momento de hablar sobre las matrices y hacerlo con algún detenimiento. No se trata de contar aquí muchas cosas acerca de ellas, sino de traer a colación algunas cuestiones que nos permitan avanzar con cierta holgura en los próximos capítulos. Nos referimos en concreto a la suma de matrices, al producto por un escalar, a la multiplicación de matrices, a la matriz inversa, cuando exista, y a las principales propiedades relacionadas con todo ello.

Poder recurrir a las matrices desde los primeros momentos, aunque ello sólo sea como apoyo, en ejemplos y en problemas, puede sernos de inestimable ayuda.

Detrás de todo esto de las matrices hay una idea básica: la de familia o sistema de elementos que dependen de dos índices. Este es el caso, por ejemplo, del horario de salidas de los trenes de una determinada línea férrea; aquí, las horas de salida de los trenes dependen de las siguientes «variables discretas» o índices: i = (número del tren) y j = (número de la estación). Si llamamos h_{ij} a la hora de salida del tren i-ésimo de la estación j-ésima, la tabla de los horarios de salida será la «matriz» cuyo elemento de lugar (i, j) es h_{ij}. En este caso, como en otros muchos, los elementos h_{ij} (horas de salida) se representan recurriendo a una «tabla de doble entrada» y se pone:

		Número de la estación(*)					
		$j=1$	$j=2$	...	j	...	$j=n$
Número del tren(**)	$i=1$	8,05	8,25	9,00	9,40	10,05	10,35
	$i=2$	9,15	9,30	9,55	10,25	10,45	11,00
	⋮	—	—	—	—	—	—
	i	—	—	—	h_{ij}	—	—
	⋮	—	—	—	—	—	—
	$i=m$	19,00	19,30	20,20	21,20	21,55	22,30

(*) También, por ejemplo, $j = 1$ puede ser «Robledal de las cabras»; $j = 2$, puede ser «Villachica del Páramo»; $j = n$ podría ser «San Veremundo de la Vera».

(**) Por ejemplo, $i = 1$ es «tren expreso»; $i = 2$ es un «tren rápido»; $i = m$ es un «tren correo».

α MATRICES; ALGEBRA DE MATRICES

Venimos recurriendo a las matrices desde las primeras páginas, pero de ellas sólo se ha dicho, por no haber necesitado de más, que son tablas rectangulares de escalares. Es obligado, pues, dar ahora las definiciones pertinentes.

3.1. PRIMERAS DEFINICIONES

[021]

Se llama *matriz* de tamaño $m \times n$, constituida por escalares de un cuerpo K (en particular, $K = \mathbb{R}$ o $K = \mathbb{C}$), a cualquier tabla rectangular A formada por $m \cdot n$ escalares, dispuestos en m filas y n columnas. Se llama elemento de lugar (i, j) o *ij de A* al escalar que está situado en la intersección de la fila i-ésima (para $i = 1, 2, ..., m$) y la columna j-ésima (para $j = 1, 2, ..., n$); si a este elemento se le llama a_{ij}, la matriz se denota poniendo:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{o} \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Abreviadamente, también se suele escribir $A = [a_{ij}]$. Cuando se quiera señalar expresamente que la matriz A tiene tamaño $m \times n$, se la denotará poniendo $A_{m\times n}$. Dos matrices de igual tamaño, $A = [a_{ij}]$ y $B = [b_{ij}]$, se dicen *matrices iguales* si tienen respectivamente iguales los elementos que ocupan los mismos lugares, esto es, si $a_{ij} = b_{ij}$ para $i = 1, 2, ..., m$ y $j = 1, 2, ..., n$.

Al conjunto de las matrices de tamaño $m \times n$ se le denota por $\mathcal{M}_{m\times n}(K)$, donde K es el cuerpo de escalares, o simplemente $\mathcal{M}_{m\times n}$. Si $m = n$, se simplifica la notación y, en lugar de $\mathcal{M}_{m\times n}$, se pone $\mathcal{M}_n$.

Si $m = n$, se dice que A es *matriz cuadrada*; si $m \neq n$, se dice que A es *rectangular*. Si para $i > j$ es $a_{ij} = 0$, se dice que A es *triangular superior*; si para $i < j$ es $a_{ij} = 0$, se dice que A es *triangular inferior*; si para $i \neq j$ es $a_{ij} = 0$, se dice que A es *matriz diagonal*. Se llama *diagonal* de A a la sucesión $a_{11}a_{22}a_{33}...$ *Matriz escalar* es una matriz diagonal que tiene iguales todos los elementos de su diagonal. Si $m = 1$, se dice que A es *matriz fila*; si $n = 1$, se dice que A es *matriz columna*. Si A es cuadrada y, para cualesquiera i y j, es $a_{ij} = a_{ji}$ o si es $a_{ij} = -a_{ji}$ se dice, respectivamente, que A es *matriz simétrica* o que A es *matriz antisimétrica*.

SOBRE LA DEFINICION DE MATRIZ

¿Qué se entiende por tabla rectangular? ¿Qué son filas y qué son columnas? Para salvar estas objeciones, se puede dar la siguiente definición de matriz:

Dados $m, n \in \mathbb{N}$ y si I y J son $I = \{1, 2, ..., m\}$ y $J = \{1, 2, ..., n\}$, se llama matriz de tamaño $m \times n$, de escalares de un cuerpo K, a toda aplicación $A : I \times J \to K$; a la imagen por A de un par (i, j) se le llama elemento de lugar (i, j) de A y le representaremos poniendo a_{ij}.

EJERCICIOS

Si $A = [a_{ij}]$ es la matriz que se da al final de este enunciado, analícese en qué

casos es: 1) triangular superior; 2) triangular inferior; 3) simétrica, esto es, tal que sus elementos a_{ij} y a_{ji} son iguales (para $i = 1, 2, 3$ y $j = 1, 2, 3$).

$$A = \begin{bmatrix} x & x+y & x-z \\ x-y & y & y+z \\ x+z-2 & z-y & z \end{bmatrix}$$

RESOLUCION

1. Han de ser $a_{21} = a_{31} = a_{32} = 0$, esto es $x - y = 0$, $x + z - 2 = 0$ y $z - y = 0$, que se verifica para $x = y = z = 1$ solamente. La solución es, pues, la matriz A_1 que figura al final de este ejercicio.
2. Ha de ser $a_{12} = a_{13} = a_{23} = 0$, esto es $x + y = 0$, $x - z = 0$ e $y + z = 0$, que se verifica para $x = -y = z$, lo que conduce a las matrices A_2 que figuran al final.
3. Ha de ser $a_{12} = a_{21}$, $a_{13} = a_{31}$ y $a_{23} = a_{32}$, esto es, $x + y = x - y$, $x - z = x + z - 2$ e $y + z = z - y$, que se verifica para $y = 0$, $z = 1$ y x cualquiera; luego las soluciones son las matrices A_3:

$$A_1 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} ; \quad A_2 = \begin{bmatrix} x & 0 & 0 \\ 2x & -x & 0 \\ 2(x-1) & 2x & x \end{bmatrix} ; \quad A_3 = \begin{bmatrix} x & x & x-1 \\ x & 0 & 1 \\ x-1 & 1 & 1 \end{bmatrix}$$

SUMA (DE MATRICES) Y PRODUCTO POR UN ESCALAR

Al dar ahora, tan pronto, las definiciones de suma de matrices y de producto de escalar por matriz, no disponemos aún de una justificación suficiente de las mismas. Por ello, nos ha parecido oportuno presentar aquí un ejemplo en el que se puede apreciar que las citadas definiciones van a ser provechosas:

EJEMPLO PRELIMINAR

Un ludópata adquiere 7 boletos de cada una de las loterías 1.ª, 2.ª y 3.ª; por cada uno de ellos paga la cantidad c_1, c_2 y c_3, respectivamente, y obtiene, en premios, las cantidades p_1, p_2 y p_3. En otra ocasión, adquiere 4 boletos y, ahora, las cantidades anteriores pasan a ser c'_1, c'_2, c'_3 y p'_1, p'_2, p'_3, respectivamente. El jugador hace balance económico (para determinar, el muy iluso, cuál de las loterías le es más propicia); los gastos y los beneficios producidos con la lotería i-ésima son $7c_i + 4c'_i$ y $7p_i + 4p'_i$. Nuestro hombre, que a pesar de todo es metódico y ordenado, realiza sus cuentas de manera conjunta y agrupa todos sus cálculos poniendo:

$$7\begin{bmatrix} c_1 & p_1 \\ c_2 & p_2 \\ c_3 & p_3 \end{bmatrix} + 4\begin{bmatrix} c'_1 & p'_1 \\ c'_2 & p'_2 \\ c'_3 & p'_3 \end{bmatrix} = \begin{bmatrix} 7c_1 + 4c'_1 & 7p_1 + 4p'_1 \\ 7c_2 + 4c'_2 & 7p_2 + 4p'_2 \\ 7c_3 + 4c'_3 & 7p_3 + 4p'_3 \end{bmatrix}$$

El ludópata de nuestro cuento acaba de inventar la suma de matrices y el producto de un escalar por una matriz.

[022]

Dadas dos matrices $A=[a_{ij}]$ y $B=[b_{ij}]$, del mismo tamaño $m \times n$ y formadas por escalares de un cuerpo cualquiera, se llama *suma* $A+B$ a la matriz $A+B=[a_{ij}+b_{ij}]$, de tamaño $m \times n$; esto es, la suma, que siempre existe, se define mediante:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

Se llama *producto* λA (λ escalar) a la matriz $\lambda A=[\lambda a_{ij}]$, de tamaño $m \times n$; esto es, el producto por escalar, que siempre existe, se define mediante:

$$\lambda \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{bmatrix}$$

Propiedades. Para cualesquiera que sean las matrices A, B y C, del mismo tamaño $m \times n$, y para cualesquiera escalares λ y μ, se verifica:

1.ª)	$(A+B)+C=A+(B+C)$	*a*)	$(\lambda+\mu)A=\lambda A+\mu A$
2.ª)	$A+O=O+A=A$ [I]	*b*)	$\lambda(A+B)=\lambda A+\lambda B$
3.ª)	$A+(-A)=(-A)+A=O$ [II]	*c*)	$\lambda(\mu A)=(\lambda\mu)A$
4.ª)	$A+B=B+A$	*d*)	$1A=A$

[I] O se llama *matriz nula*, de tamaño $m \times n$, y es aquella cuyos elementos son todos iguales al escalar nulo.

[II] $-A=[-a_{ij}]$ se llama *matriz opuesta* de $A=[a_{ij}]$ y es aquella, de tamaño $m \times n$, cuyos elementos son los escalares opuestos de los respectivos elementos de A

EL ESPACIO VECTORIAL $\mathcal{M}_{m\times n}$

De aquí a poco, cuando en un capítulo próximo hablemos de la estructura de «espacio vectorial», tendremos ocasión de comprobar que, allí, se arrancará de un conjunto donde se definen una suma y un producto por escalar que han de satisfacer a unos axiomas, que no son sino las propiedades 1.ª, 2.ª, 3.ª, 4.ª, *a*, *b*, *c* y *d* que, aquí, se estudian para las matrices.

COMPROBACION

Todas estas propiedades se demuestran con gran facilidad.Para ello, bastará con recurrir a que, como ya sabemos (véase la definición de cuerpo), las citadas propiedades son ciertas para los escalares, esto es, se verifican si se sustituyen las matrices A, B y C por escalares a, b y c (o a_{ij}, b_{ij} y c_{ij}). Así, por ejemplo, la propiedad $\lambda(A+B)=\lambda A+\lambda B$ se satisface como consecuencia de ser $\lambda(a_{ij}+b_{ij})=\lambda a_{ij}+\lambda b_{ij}$ (para $i=1, 2, ..., m$ y $j=1, 2, ..., n$).

EJERCICIO

Hallar dos matrices X e Y, de tamaño 2×3, tales que

$$\left.\begin{aligned} 3X+Y=A\\ 4X+2Y=B\end{aligned}\right\}$$

siendo,

$$A=\begin{bmatrix}3/2 & 1 & 0\\ 2 & 1/2 & 5\end{bmatrix} \quad \text{y} \quad B=\begin{bmatrix}3 & 4 & 2\\ 2 & 1 & 8\end{bmatrix}$$

RESOLUCION

Operando de acuerdo con las anteriores propiedades, esto es, igual que si X, Y, A y B fuesen escalares, el sistema dado es equivalente a cada uno de los siguientes:

$$\left.\begin{aligned} 6X+2Y=2A\\ 4X+2Y=B\end{aligned}\right\} \quad \left.\begin{aligned} -3X+A=Y\\ 2X=2A-B\end{aligned}\right\} \quad \left.\begin{aligned} X=A-\tfrac{1}{2}B\\ Y=-2A+\tfrac{3}{2}B\end{aligned}\right\}$$

Por tanto, X e Y han de valer:

$$X=A-\tfrac{1}{2}B=\begin{bmatrix}0 & -1 & -1\\ 1 & 0 & 1\end{bmatrix} \;;\; Y=-2A+\tfrac{3}{2}B=\begin{bmatrix}-3/2 & 0 & 1\\ -3 & -1/2 & -6\end{bmatrix}$$

3.2. PRODUCTO DE MATRICES

Es seguro que, a estas alturas, todos conocemos ya algo, aunque sea poco, sobre el producto de matrices y que, por ello, es probable que no nos resulte chocante, cuando no estrafalaria, la manera de definir dicho producto. Pero, además, la multiplicación de matrices no se define, en general, cuando éstas tienen el mismo tamaño, como podría esperarse, sino que para que un producto AB tenga sentido, va a ser necesario que A tenga tantas columnas como filas tiene B. Y es que esta definición es la que es, y no es otra, para que sea de utilidad en el momento de componer aplicaciones lineales, cosa ésta que queda para más adelante.

Como aquí se ha optado por presentar a las matrices lo antes posible, para que sirva de herramienta desde el primer momento, nos vemos obligados a dar una pequeña justificación de la referida definición de producto, lo que se intenta con el siguiente ejemplo.

EJEMPLO PRELIMINAR

Supongamos que en un centro de enseñanza que se llama «Instituto Superior de Futurología» se va a seleccionar a tres becarios, de entre los cinco aspirantes que abajo se relacionan, y que ello se hace atendiendo a las calificaciones que hayan obtenido en determinadas materias. Los aspirantes y sus calificaciones son los que figuran en la siguiente tabla o matriz, que llamaremos *A*:

A — Calificaciones obtenidas por los aspirantes

Aspirantes	Materias: Astrología	Cartomancia	Auguración	Aruspicina
Sigisberta	a_{11} 7	a_{12} 5	a_{13} 8	a_{14} 10
Prócolo	a_{21} 5	a_{22} 9	a_{23} 7	a_{24} 8
Geroncia	a_{31} 10	a_{32} 7	a_{33} 6	a_{34} 8
Báyulo	a_{41} 9	a_{42} 6	a_{43} 10	a_{44} 7
Juventino	a_{51} 8	a_{52} 7	a_{53} 9	a_{54} 8

Para la concesión de las becas se han establecido unos baremos que, según al departamento al que vaya a ir destinado el becario (que llamaremos departamentos 1.°, 2.° y 3.°), dan distinto «peso» a cada una de las calificaciones obtenidas por los aspirantes. Estos «pesos» son los que se señalan en la siguiente matriz, que llamaremos *B*:

B — «Pesos» del baremo de selección

Materia	Departamento: 1.°	2.°	3.°
1.ª) Astrología	b_{11} 4	b_{12} 1	b_{13} 1,5
2.ª) Cartomancia	b_{21} 1	b_{22} 4,5	b_{23} 1
3.ª) Auguración	b_{31} 3.5	b_{32} 1,5	b_{33} 2,5
4.ª) Aruspicina	b_{41} 1.5	b_{42} 3	b_{43} 5
Totales	10	10	10

Estos baremos funcionan linealmente, esto es, para cada departamento, la puntuación total que obtiene un aspirante es la suma de los productos de sus notas por los respectivos «pesos» que se les asigna en el departamento en cuestión. Así, el aspirante Báyulo (aspirante número 4) para el departamento 3.° obtiene la puntuación c_{43} siguiente:

$$c_{43} = 9 \times 1{,}5 + 6 \times 1 + 10 \times 2{,}5 + 7 \times 5 = 79{,}5$$

o sea,

$$c_{43} = a_{41}b_{13} + a_{42}b_{23} + a_{43}b_{33} + a_{44}b_{43} = \sum_{h=1}^{4} a_{4h}b_{h3}$$

Procediendo de igual modo con los demás aspirantes y departamentos, se obtienen todas las puntuaciones; al disponer éstas en la correspondiente tabla de doble entrada, se obtiene la siguiente matriz, que llamaremos C:

C	Puntuación total de los aspirantes en cada departamento	Departamentos		
		1.°	2.°	3.°
Aspirantes	Sigisberta	c_{11} 76	c_{12} 71,5	c_{13} 87,5
	Prócolo	c_{21} 65,5	c_{22} 80	c_{23} 74
	Geroncia	c_{31} 80	c_{32} 74,5	c_{33} 77
	Báyulo	c_{41} 87,5	c_{42} 72	c_{43} 79,5
	Juventino	c_{51} 82,5	c_{52} 77	c_{53} 81,5

NOTA: Los aspirantes elegidos (los de mayor puntuación) son: para el primer departamento, Báyulo; para el segundo, Prócolo; para el tercero, Sigisberta.

Donde el elemento de lugar (ij) de la matriz C es la suma de los productos de los elementos de la fila i-ésima de A por los correspondientes elementos de la columna j-ésima de B; esto es:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + a_{i4}b_{4j} = \sum_{h=1}^{4} a_{ih}b_{hj}$$

Se dice que esta matriz C es el producto de la matriz A por la matriz B; este producto ha podido realizarse debido a que el número de columnas de A es igual al número de filas de B. Se pone:

$$C = AB \quad \text{o} \quad [c_{ij}] = [a_{ih}][b_{hj}]$$

DEFINICION DE PRODUCTO DE MATRICES

[023]

Dadas las matrices $A = [a_{ih}]$ cuyo tamaño es $m \times p$ y $B = [b_{hj}]$ de tamaño $p \times n$ (nótese que B tiene tantas filas, p, como columnas tiene A), constituidas ambas por escalones de un mismo cuerpo, se llama *producto AB* o $A \cdot B$, de A por B, a la matriz $C = [c_{ij}]$ de tamaño $m \times n$ cuyo elemento de lugar (ij) es:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{h=1}^{p} a_{ih}b_{hj}$$

$$\begin{matrix} i \rightarrow \\ \\ \\ \\ \end{matrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{ip} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \vdots & b_{1j} & \vdots & b_{1n} \\ b_{21} & \vdots & b_{2j} & \vdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{p1} & \vdots & b_{pj} & \vdots & b_{pn} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ & c_{ij} & \\ c_{m1} & \cdots & c_{mn} \end{bmatrix} \begin{matrix} \\ \\ \leftarrow i \\ \\ \end{matrix}$$

$$\begin{matrix} \\ \\ \\ \\ \uparrow \\ j \end{matrix} \qquad \begin{matrix} \\ \\ \\ \\ \uparrow \\ j \end{matrix}$$

Así, por ejemplo, al multiplicar las siguientes matrices, de tamaños 2×4 la primera y 4×3 la segunda, se obtiene la matriz 2×3 que se indica:

$$\begin{bmatrix} 2 & -1 & 5 & -3 \\ 4 & 2 & -3 & -5 \end{bmatrix} \begin{bmatrix} 1 & -2 & 3 \\ 3 & 6 & -1 \\ 2 & 7 & -2 \\ 4 & 1 & 5 \end{bmatrix} =$$

$$\begin{matrix} 2\times 1 + (-1)\times 3 + 5\times 2 + (-3)\times 4 \\ 2\times(-2) + (-1)\times 6 + 5\times 7 + (-3)\times 1 \\ 2\times 3 + (-1)\times(-1) + 5\times(-2) + (-3)\times 5 \end{matrix}$$

$$= \begin{bmatrix} -3 & 22 & -18 \\ -16 & -22 & -9 \end{bmatrix}$$

$$\begin{matrix} 4\times 1 + 2\times 3 + (-3)\times 2 + (-5)\times 4 \\ 4\times(-2) + 2\times 6 + (-3)\times 7 + (-5)\times 1 \\ 4\times 3 + 2\times(-1) + (-3)\times(-2) + (-5)\times 5 \end{matrix}$$

CASOS PARTICULARES

Concretamos aquí el valor del producto AB para algunos casos de especial interés:

• Si A es matriz fila $(1 \times p)$ y B es cualquiera $(p \times n)$, entonces:

$$[a_1 \quad a_2 \quad \cdots \quad a_p]\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{bmatrix} = [c_1 \quad c_2 \quad \cdots \quad c_n]$$

$$\text{donde} \quad c_i = \sum_{h=1}^{p} a_h b_{hi} \quad (i = 1, 2, ..., n)$$

• Si A es cualquiera $(m \times p)$ y B es matriz columna $(p \times 1)$, entonces:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_p \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ \vdots \\ c_m \end{bmatrix}$$

$$\text{donde} \quad c_i = \sum_{h=1}^{p} a_{ih} b_h \quad (i = 1, 2, ..., m)$$

• Si A es una matriz fila $(1 \times p)$ y B es una matriz columna $(p \times 1)$, entonces:

$$[a_1 \quad a_2 \quad \cdots \quad a_p]\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix} = [c]$$

$$\text{donde} \quad c = \sum_{h=1}^{p} a_h b_h$$

• Si A es matriz columna $(m \times 1)$ y B es matriz fila $(1 \times n)$, entonces:

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_m \end{bmatrix}[b_1 \quad b_2 \quad \cdots \quad b_n] = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix}$$

$$\text{donde} \quad c_{ij} = a_i b_j \quad (i = 1, 2, ... m; j = 1, 2, ..., n)$$

• Si una de las matrices que se multiplican es una matriz escalar E (véase [021]), entonces:

$$AE = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} e & 0 & \cdots & 0 \\ 0 & e & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & e \end{bmatrix} = \begin{bmatrix} a_{11}e & a_{12}e & \cdots & a_{1p}e \\ a_{21}e & a_{22}e & \cdots & a_{2p}e \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}e & a_{m2}e & \cdots & a_{mp}e \end{bmatrix} = eA$$

$$EB = \begin{bmatrix} e & 0 & \cdots & 0 \\ 0 & e & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & e \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{bmatrix} = \begin{bmatrix} eb_{11} & eb_{12} & \cdots & eb_{1n} \\ eb_{21} & eb_{22} & \cdots & eb_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ ebp_1 & ebp_2 & \cdots & eb_{pn} \end{bmatrix} = eB$$

EJERCICIO

Estudiar si existe alguna matriz real A de tamaño 3×2 tal que $AA^t = I$, donde A^t (traspuesta de A) se obtiene de cambiar filas por columnas en A e I es la matriz unidad (de tamaño 3×3).

RESOLUCION

Hay que analizar si existen algunos números reales a, a', b, b', c y c' tales que:

$$\begin{bmatrix} a & a' \\ b & b' \\ c & c' \end{bmatrix} \begin{bmatrix} a & b & c \\ a' & b' & c' \end{bmatrix} \equiv \begin{bmatrix} a^2+a'^2 & ab+a'b' & ac+a'c' \\ ab+a'b' & b^2+b'^2 & bc+b'c' \\ ac+a'c' & bc+b'c' & c^2+c'^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Como $a^2+a'^2$, $b^2+b'^2$ y $c^2+c'^2$ han de valer 1, han de existir θ_1, θ_2 y θ_3 (ángulos entre 0 y 2π) tales que

$$a = \cos\theta_1, \quad b = \cos\theta_2, \quad c = \cos\theta_3$$
$$a' = \operatorname{sen}\theta_1, \quad b' = \operatorname{sen}\theta_2, \quad c' = \operatorname{sen}\theta_3$$

Por otra parte, como $ab + a'b' = 0$, resulta que $\cos(\theta_1 - \theta_2) = 0$, o sea $\theta_1 - \theta_2 = \pm\pi/2, \pm 3\pi/2$; análogamente $\theta_1 - \theta_3$ y $\theta_2 - \theta_3$ valdrían $\pm\pi/2$ ó $\pm 3\pi/2$. Este resultado es contradictorio ya que, de ser así, como se puede poner $\theta_1 - \theta_3 = (\theta_1 - \theta_2) + (\theta_2 - \theta_3)$ nos encontraríamos con que $\theta_1 - \theta_3$ debería valer 0, $\pm\pi$ ó $\pm 2\pi$ y no $\pm\pi/2$ ó $\pm 3\pi/2$ como se acaba de decir. Por tanto, el problema propuesto carece de solución.

PROPIEDADES DEL PRODUCTO

Así como para la suma de matrices y para el producto por un escalar, se verificaban las mismas leyes que rigen para los escalares, para el producto la cosa es distinta; ahora, aunque algunas leyes siguen siendo ciertas, otras ya no son aquí de aplicación.

[024]

Para cualesquiera que sean las siguientes matrices (*), se verifican las propiedades:

1. $(AB)C = A(BC)$ (asociativa).
2. $\left.\begin{array}{l} A(B+B') = AB + AB' \\ (A+A')B = AB + A'B \end{array}\right\}$ (distributivas).
3. $AI = A$ e $IB = B$, donde I es la siguiente matriz cuadrada, llamada *matriz unidad*:

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

4. $AB \neq BA$ (en general); esto es, el producto no es conmutativo.
5. Puede ser $AB = O$ sin que sean $A = O$ o $B = O$ (cuando sea $AB = O$, $A \neq O$ y $B \neq O$, se dirá que A y B son *divisores de cero*).
6. El conjunto $\mathcal{M}_n$ de las matrices cuadradas de cierto tamaño $n \times n$ (donde n es cualquiera, pero fijo), con las operaciones suma y producto de matrices, es un anillo unitario no conmutativo y con divisores de cero.

(*) Se exige, obviamente, que tengan sentido todos los productos que aquí intervienen. Los tamaños de las matrices deberán ser, pues, los: A y A' de tamaño $m \times p$; B y B' de tamaño $p \times n$; C de tamaño $n \times q$; I cuadrada de tamaño $p \times p$.

COMPROBACIONES

1. Llamando $A = [a_{ih}]$, $B = [b_{hj}]$ y $C = [c_{jk}]$, se tiene:

$$(AB)C = \left[\sum_h a_{ih} b_{hj}\right][c_{jk}] = \left[\sum_j \left(\sum_h a_{ih} b_{hj}\right) c_{jk}\right] = \left[\sum_{jh} a_{ih} b_{hj} c_{jk}\right]$$

$$A(BC) = [a_{ih}]\left[\sum_j b_{hj} c_{jk}\right] = \left[\sum_h a_{ih}\left(\sum_j b_{hj} c_{jk}\right)\right] = \left[\sum_{hj} a_{ih} b_{hj} c_{jk}\right]$$

Estos dos resultados, esto es, los elementos de lugar (ik) de $(AB)C$ y de

$A(BC)$, son evidentemente iguales para cualesquiera que sean i y k, luego $(AB)C = (AB)C$.

2. Llamando $A = [a_{ih}]$, $B = [b_{hj}]$ y $B' = [b'_{hj}]$, se tiene:

$$A(B+B') = [a_{ih}][b_{hj} + b'_{hj}] = \left[\sum_h a_{ih}(b_{hj} + b'_{hj})\right] = \left[\sum_h a_{ih}b_{hj} + \sum_h a_{ih}b'_{hj}\right] =$$

$$= \left[\sum_h a_{ih}b_{hj}\right] + \left[\sum_h a_{ih}b'_{hj}\right] = [a_{ih}][b_{hj}] + [a_{ih}][b'_{hj}] = AB + AB'$$

Análogamente si comprueba que es $(A + A')B = AB + A'B$.

3. Las igualdades $AI = A$ e $IB = B$ se obtienen de hacer $e = 1$ en el último de los «casos particulares» del apartado anterior, esto es, tomando como caso particular de matriz escalar E a la matriz unidad I.

4. Para que existan los dos productos AB y BA, es necesario que los tamaños de A y de B sean $m \times n$ y $n \times m$ (para algunos m y n); en tal caso, AB es de tamaño $m \times m$ y BA lo es de tamaño $n \times n$. Para que AB y BA existan y sean de igual tamaño hace falta, pues. que sea $m = n$, es decir, A y B deberán ser cuadradas y de igual tamaño $(n \times n)$. Aún en este caso, AB y BA serán, generalmente, distintas(*); así, por ejemplo; para las siguientes matrices A y B de tamaño 3×3, es $AB \neq BA$:

$$A = \begin{bmatrix} 1 & 2 & 5 \\ -2 & 1 & 3 \\ 1 & 0 & -1 \end{bmatrix} \quad ; \quad B = \begin{bmatrix} -3 & 0 & 2 \\ 4 & 3 & -3 \\ -1 & -1 & 1 \end{bmatrix}$$

$$AB = \begin{bmatrix} 0 & 1 & 1 \\ 7 & 0 & -4 \\ -2 & 1 & 1 \end{bmatrix} \quad ; \quad BA = \begin{bmatrix} -1 & -6 & -17 \\ -5 & 11 & 32 \\ 2 & -3 & -9 \end{bmatrix}$$

5. Comprobemos que existen divisores de cero, dando algunos ejemplos:

$$\begin{bmatrix} 1 & 0 & -2 & 1 \\ 3 & -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & -3 \\ 2 & 1 & -7 \\ -1 & 0 & -1 \\ -4 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = O$$

$$\begin{bmatrix} 1 & -3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 3 & 9 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = O$$

6. Para este conjunto $\mathcal{M}_n$, de las matrices cuadradas $n \times n$, además de la suma, también el producto es una operación interna, esto es, dos matrices $n \times n$ siempre se pueden sumar y multiplicar y los resultados son también matrices del mismo tamaño $n \times n$. Pues bien, $\mathcal{M}_n$ con estas dos operaciones es un anillo ya que: *a*) por verificarse las propiedades 1.ª, 2.ª, 3.ª y 4.ª de la suma (véase [022]), $\mathcal{M}_n$ es un grupo abeliano; y *b*) se verifica también la asociatividad del producto y las distributivas (que son las propiedades 1.ª

(*) Si A y B son tales que $AB = BA$, se dice que A y B conmutan o permutan entre sí.

y 2.ª recién comprobadas). Este anillo es unitario, no conmutativo y con divisores de cero, según se acaba de comprobar en las anteriores propiedades 3.ª, 4.ª y 5.ª.

3.3. TRASPUESTA DE UNA MATRIZ

A cada matriz A se le va a asociar otra matriz A^t, que se llama traspuesta de aquella; A^t es el resultado de cambiar, en A, las filas por las columnas. Aquí, al tiempo que se define la trasposición de matrices, también se consideran sus principales propiedades.

[025]

Dada una matriz A, de tamaño $m \times n$, se llama *traspuesta* de A a la matriz A^t, de tamaño $n \times m$, que tiene por elemento de lugar (ij) al elemento de lugar (ji) de A, para $i = 1,2, ..., n$ y $j = 1, 2, ..., m$. La correspondencia $A \to A^t$ se llama *trasposición* de matrices.

Propiedades. Para cualesquiera matrices A y B, se verifica:

1.ª $(A^t)^t = A$ (la trasposición es involutiva).
2.ª $(A + B)^t = A^t + B^t$ (A y B del mismo tamaño).
3.ª $(\lambda A)^t = \lambda A^t$ (λ escalar cualquiera).
4.ª $(AB)^t = B^t A^t$ (A de tamaño $m \times p$; B de tamaño $p \times m$).

Si A es una matriz cuadrada, decir que A es simétrica equivale a decir que $A = A^t$; decir que A es antisimétrica equivale a decir que $A = -A^t$. Para cualquiera que sea la matriz cuadrada M, la matriz $M + M^t$ es simétrica y la $M - M^t$ es antisimétrica.

COMPROBACION

Con excepción de la propiedad 4.ª, que se prueba a continuación, el resto de las afirmaciones del enunciado se comprueban trivialmente. Para la 4.ª, llamando a_{rs}, a'_{rs}, b_{rs} y b'_{rs} a los elementos de lugar (rs) de A, A^t, B y B^t, respectivamente, se tiene:

$$[\text{elemento } (ij) \text{ de } (AB)^t] = [\text{elemento } (ji) \text{ de } AB] = \sum_h a_{jh} b_{hi}$$

$$[\text{elemento } (ij) \text{ de } B^t A^t] = \sum_h b'_{ih} a'_{hj} = \sum_h b_{hi} a_{jh} = \sum_h a_{jh} b_{hi}$$

Como $(AB)^t$ y $B^t A^t$ han resultado tener los mismos elementos, han de ser iguales.

EJERCICIO

Compruébese que cualquier matriz cuadrada M se puede expresar como suma de dos matrices que son una simétrica, S, y otra antisimétrica, A. Compruébese también que tal descomposición es única.

RESOLUCION

La descomposición $M = S + A$ se consigue tomando

$$S = \tfrac{1}{2}(M + M^t) \quad \text{y} \quad A = \tfrac{1}{2}(M - M^t)$$

donde S y A son, según sabemos, simétrica y antisimétrica, respectivamente. Supongamos que se dispone de otra descomposición $M = S' + A'$; de ser así, resultaría que

$$S + A = S' + A' \quad \text{luego} \quad S - S' = A' - A$$

Nótese que $S - S'$ es simétrica y $A' - A$ es antisimétrica. Llamando s_{ij} y a_{ij} a los elementos de lugar (ij) de $S - S'$ y de $A' - A$, como estas matrices son iguales, ha de ser $s_{ij} = a_{ij}$ y $s_{ij} = a_{ji}$; ahora bien, como $s_{ji} = s_{ij}$ y $a_{ji} = a_{ij}$, resulta que

$$s_{ij} = a_{ij} \quad \text{y} \quad s_{ij} = -a_{ij}$$

(para cualesquiera i y j). De aquí se deduce, pues, que $s_{ij} = a_{ij} = 0$ para cualesquiera i y j, esto es, que $S - S' = O$ y $A - A' = O$. o sea $S = S'$ y $A = A'$ que prueba la unicidad de la descomposición.

Lo que aquí se ha hecho no ha sido otra cosa que expresar una matriz cualquiera $M = [m_{ij}]$ en la forma:

$$[m_{ij}] = \tfrac{1}{2}[m_{ij} + m_{ji}] + \tfrac{1}{2}[m_{ij} - m_{ji}]$$

Así, por ejemplo:

$$\begin{bmatrix} 2 & 4 & 5 \\ 3 & -1 & 6 \\ 0 & -3 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 7/2 & 5/2 \\ 7/2 & -1 & 3/2 \\ 5/2 & 3/2 & 7 \end{bmatrix} + \begin{bmatrix} 0 & 1/2 & 5/2 \\ -1/2 & 0 & 9/2 \\ -5/2 & -9/2 & 0 \end{bmatrix}$$

3.4. RELACION ENTRE LAS OPERACIONES ELEMENTALES Y EL PRODUCTO

Empecemos recordando que las operaciones elementales (véase [016]), sobre un sistema de vectores (que aquí serán las filas o las columnas de una matriz), son de uno de los siguiente tipos:

- Tipo $[i \leftrightarrow j]$ consiste en intercambiar los vectores i-ésimo y j-ésimo.
- Tipo $[i \rightarrow \lambda i]$ consistente en multiplicar el vector i-ésimo por el escalar $\lambda \neq 0$.
- Tipo $[i \rightarrow i + j]$ consiste en sumar el vector j-ésimo al vector i-ésimo.
- La aplicación reiterada de cualesquiera operaciones de los tipos anteriores.

Vamos ahora a comprobar que las operaciones elementales, en una matriz A, se pueden realizar, también, multiplicando a A por unas matrices adecuadas; si las operaciones elementales son en filas, hay que multiplicar por la izquierda (premultiplicar), de A, y si lo son en columnas, hay que multiplicar por la derecha (posmultiplicar):

[026]

Sea A una matriz de tamaño $m \times n$ dada. Se verifica que:

1. Dada una operación elemental E^f o E^c, sobre filas o sobre columnas respectivamente, al aplicársela a la matriz A se obtiene una matriz $E^f(A)$ o $E^c(A)$, que vale:

$$E^f(A) = FA, \quad \text{donde} \quad F = E^f(I_m), \quad I_m = \text{unidad } m \times m$$
$$E^c(A) = AC, \quad \text{donde} \quad C = E^c(I_n), \quad I_n = \text{unidad } n \times n$$

2. Si E es la operación elemental que consiste en aplicar una operación E^f a las filas y una operación E^c a las columnas, al realizar una tras otra (en cualquier orden), la matriz A se transforma en la matriz $E(A)$, que vale:

$$E(A) = FAC, \quad \text{donde} \quad \begin{cases} F = E^f(I_m) \\ C = E^c(I_n) \end{cases}$$

3. Si la matriz dada A tiene rango r, entonces hay ciertas operaciones E^f y E^c (sobre filas y columnas, respectivamente) tales que:

$$FAC = \left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right], \quad \text{donde} \quad \begin{cases} F = E^f(I_m) \\ C = E^c(I_n) \\ I_r = \text{unidad } r \times r \end{cases}$$

NOTA: Recuérdese que a esta última matriz se le llamó (en [020]) matriz canónica de equivalencia de la matriz dada A.

DEMOSTRACION

Se suelen llamar *matrices elementales* a las matrices cuadradas, de cualquier tamaño $p \times p$, que se obtiene de aplicar a la matriz unidad, de tamaño $p \times p$, las operaciones elementales de los tres tipos básicos, esto es, las $\dot{E} = [i \leftrightarrow j]$,

$\ddot{E} = [i \rightarrow \lambda I]$ con $\lambda \neq 0$ y $\dddot{E} = [i \rightarrow i + j]$. Las matrices elementales son, pues, las siguientes:

$$\dot{E}^{f}(I) = \begin{bmatrix} 1 & & & & \\ & 0 & \cdots & 1 & \leftarrow i \\ & \vdots & & \vdots & \\ & 1 & \cdots & 0 & \leftarrow j \\ & & & & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \end{matrix}$$
$$\begin{matrix} & \uparrow & & \uparrow & \\ & i & & j & \end{matrix}$$

Matriz cuadrada cuyas únicos elementos no nulos son: 1) los elementos de la diagonal, que valen 1, excepto los dos ceros que se señalan; y 2) los elementos de lugares ij y ji que valen 1. Aquí $\dot{E}$ es la operación $\dot{E} = [i \leftrightarrow j]$.

$$\ddot{E}^{c}(I) = \begin{bmatrix} 1 & \vdots & & \\ \cdots & \lambda & \cdots\cdots & \leftarrow i \\ & \vdots & & \\ & \vdots & & \\ & \vdots & & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \end{matrix}$$
$$\begin{matrix} & \uparrow & \\ & i & \end{matrix}$$

Matriz cuadrada que tiene nulos todos los elementos no situados en la diagonal; los de la diagonal valen 1 excepto el de lugar ii, que se señala, que vale $\lambda \neq 0$. Aquí $\ddot{E}$ es la operación $\ddot{E} = [i \rightarrow \lambda i]$.

$$\dddot{E}^{f}(I) = \begin{bmatrix} 1 & & & & \\ & 1 & \cdots & 1 & \leftarrow i \\ & & & \vdots & \\ & & & 1 & \leftarrow j \\ & & & & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \end{matrix}$$
$$\begin{matrix} & \uparrow & & \uparrow & \\ & i & & j & \end{matrix}$$

Matriz cuadrada que tiene todos los elementos de su diagonal valiendo 1, su elemento de lugar ij también vale 1 y todos sus demás elementos son nulos. Aquí $\dddot{E}$ es la operación $\dddot{E} = [i \rightarrow i + j]$.

También son matrices elementales las $\dot{E}^{c}(I)$, $\ddot{E}^{c}(I)$ y $\dddot{E}^{c}(I)$, que son las traspuestas de las $\dot{E}^{f}(I)$, $\ddot{E}^{f}(I)$ y $\dddot{E}^{f}(I)$, respectivamente.

1. Sólo vamos a ocuparnos de las operaciones elementales sobre las filas; para las operaciones sobre las columnas se puede razonar de manera análoga. Empezaremos abordando los casos en los que la operación E es una de las $\dot{E}$, $\ddot{E}$ y $\dddot{E}$. Para probar cada uno de ellos. no hay mas que premultiplicar

A por cada una de las matrices elementales $\dot{E}^f(I)$, $\ddot{E}^f(I)$ y $\dddot{E}^f(I)$ y comprobar que el resultado es igual a $\dot{E}^f(A)$, $\ddot{E}^f(A)$ y $\dddot{E}^f(A)$, respectivamente; así ocurre, en efecto, ya que obviamente:

$\dot{E}^f(I)A$ y $\dot{E}^f(A)$ con ambas iguales a:	$\ddot{E}^f(I)A$ y $\ddot{E}^f(A)$ son ambas iguales a:	$\dddot{E}^f(I)A$ y $\dddot{E}^f(A)$ son ambas iguales a:
$i\rightarrow$ [fila j de A]	$i\rightarrow$ λ [fila i de A]	$i\rightarrow$ [fila i + fila j]
$j\rightarrow$ [fila i de A]		

(sólo se señalan las filas que son diferentes de las de A; las que coinciden con las de A se denotan poniendo simplemente puntos suspensivos).

Recurriendo a que la propiedad ha resultado ser cierta para las tres operaciones elementales básicas, se puede comprobar que también lo es para cualquier otra operación elemental E, pues ésta es composición de aquéllas. La comprobación de esto se puede reducir al caso en el que E es la composición de dos operaciones E_1 y E_2 (por ejemplo, $\dot{E}$ y $\ddot{E}$) para las que ya se cumple la propiedad; dicha comprobación es evidente (para simplificar la notación, se ha puesto F, F_1 y F_2 en lugar de $E^f(I)$, $E_1^f(I)$ y $E_2^f(I)$, respectivamente):

$$E^f(A) = E_2^f(E_1^f(A)) = F_2 \cdot E_1^f(A) = F_2 \cdot F_1 \cdot A = F_2 \cdot F_1 \cdot I_m \cdot A =$$
$$= (F_2 \cdot F_1 \cdot I_m)A = E^f(I)A = FA$$

2. Esta propiedad es consecuencia inmediata de la anterior, ya que:

$$E(A) = E^c(E^f(A)) = E^f(A)C = FAC \quad \text{(si primero se aplica } E^f\text{)}$$
$$E(A) = E^f(E^c(A)) = FE^c(A) = FAC \quad \text{(si primero se aplica } E^c\text{)}$$

3. Sabemos (véase [019]) que hay unas ciertas operaciones elementales, que ahora llamaremos E^f y E^c, tales que aplicando una tras otra (E^f a filas y E^c a columnas) transforman la matriz A en la matriz canónica de equivalencias, que llamaremos C_r. Recurriendo, pues, a la anterior propiedad 2, se tiene que

$$C_r = E^c(E^f(A)) = E^f(A)C = FAC \quad \text{(si primero se aplica } E^f\text{)}$$
$$C_r = E^f(E^c(A)) = FE^c(A) = FAC \quad \text{(si primero se aplica } E^c\text{)}$$

OBSERVACION

A la vista de cuanto venimos diciendo, resulta evidente que: Si E es la operación elemental que consiste en aplicar las operaciones en filas E_1^f, E_2^f, ..., E_h^f (en este orden) y las operaciones en columnas E_1^c, E_2^c, ..., E_k^c (en este orden), entonces la transformada $E(A)$, de una matriz A (de tamaño $m \times n$), es igual a:

$$E(A) = FAC$$

donde:

$$\begin{cases} F = F_h \ldots F_2 F_1 & \text{siendo } F_i = E_i^f(I) \quad \text{para} \quad i = 1,\ 2,\ \ldots,\ h \\ C = C_1 C_2 \ldots C_k & \text{siendo } C_j = E_j^c(I) \quad \text{para} \quad j = 1,\ 2,\ \ldots,\ k \end{cases}$$

EJEMPLO

Realicemos operaciones elementales en la siguiente matriz A, de tamaño 3×4, primero en las filas y luego en las columnas, hasta llegar a su matriz canónica de equivalencia C_r. Si, al tiempo, realizamos las mismas operaciones elementales en la correspondiente matriz unidad (cuando las operaciones son en filas se realizan en $I_{3\times3}$; cuando son en columnas se realizan en $I_{4\times4}$), obtendremos también dos matrices F y C, de tamaños 3×3 y 4×4, que permiten poner $FAC = C_r$. En concreto, con las operaciones elementales que se indican delante de cada una de las siguientes etapas, se va obteniendo sucesivamente:

$$A = \begin{bmatrix} 1 & 3 & 2 & -1 \\ 2 & 4 & 2 & 1 \\ 4 & 6 & 2 & 5 \end{bmatrix} \ ; \ I_{3\times3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \ ; \ I_{4\times4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

1º $\left.\begin{array}{r} 2ª\,f \to 2ª\,f - 2 \cdot 1ª\,f \\ 3ª\,f \to 3ª\,f - 4 \cdot 1ª\,f \\ (\text{sobre las } I_{3\times3} \text{ y } A) \end{array}\right\} \to \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 3 & 2 & -1 \\ -2 & 1 & 0 & 0 & -2 & -2 & 3 \\ -4 & 0 & 1 & 0 & -6 & -6 & 9 \end{array}\right] = [F' \,|\, A']$

2º $\left.\begin{array}{r} 3ª\,f \to 3ª\,f - 3 \cdot 2ª\,f \\ (\text{sobre las } F' \text{ y } A') \end{array}\right\} \to \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 3 & 2 & -1 \\ -2 & 1 & 0 & 0 & -2 & -2 & 3 \\ 2 & -3 & 1 & 0 & 0 & 0 & 0 \end{array}\right] = [F'' \,|\, A'']$

3º $\left.\begin{array}{r} 2ª\,f \to (-1/2)2ª\,f \\ (\text{sobre las } F'' \text{ y } A'') \end{array}\right\} \to \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 3 & 2 & -1 \\ 1 & -1/2 & 0 & 0 & 1 & 1 & -3/2 \\ 2 & -3 & 1 & 0 & 0 & 0 & 0 \end{array}\right] = [F''' \,|\, A''']$

4º

$$\begin{array}{l} 2ª\,c \to 2ª\,c - 3 \cdot 1.ª\,c \\ 3ª\,c \to 3ª\,c - 2 \cdot 1ª\,c \\ 4ª\,c \to 4ª\,c + 1ª\,c \\ (\text{sobre } A''' \text{ e } I_{4\times4}) \end{array} \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -3/2 \\ 0 & 0 & 0 & 0 \\ \hline 1 & -3 & -2 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c} A^{IV} \\ \hline C' \end{array}\right]$$

5°

$3^a c \rightarrow 3^a c - 2^a c$

$4^a c \rightarrow 4^a c + (3/2)2^a c$

(sobre las A^{IV} y C')

$$\left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & -3 & 1 & -7/2 \\ 0 & 1 & -1 & 3/2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c} A^V \\ \hline C'' \end{array}\right]$$

Todo esto, nos permite poner $FAC = C_r$, donde $F = F'''$, $C = C''$ y $C_r = A^V$; esto es:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & -1/2 & 0 \\ 2 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 & -1 \\ 2 & 4 & 2 & 1 \\ 4 & 6 & 2 & 5 \end{bmatrix} \begin{bmatrix} 1 & -3 & 1 & -7/2 \\ 0 & 1 & -1 & 3/2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right]$$

3.5. MULTIPLICACION DE MATRICES POR BLOQUES

SUBMATRICES Y BLOQUES

Dada una matriz A, se llama *submatriz* de A que definen los índices de filas $i_1, i_2, ..., i_p$ y los índices de columnas $j_1, j_2, ..., j_q$ a la matriz, de tamaño $p \times q$, cuyo elemento de lugar (h, k) es el elemento de lugar (i_h, j_k) de A. Los elementos de la submatriz son aquellos en los que se cruzan las filas y columnas elegidas. Si los índices (de filas y de columnas) son consecutivos, la submatriz se llama *bloque* o *caja*.

DESCOMPOSICION EN BLOQUES

Dada una matriz A, de tamaño $m \times n$, considérese una sucesión creciente de índices de filas $0 < m_1 < m_2 < \cdots < m_p = m$ y una sucesión creciente de índices de columnas $0 < n_1 < n_2 < \cdots < n_q = n$. Sea A_{ij} el bloque de A que definen las filas comprendidas entre las $m_{i-1} + 1$ y m_i (ambas inclusive) y las columnas comprendidas entre las $n_{j-1} + 1$ y n_j (ambas inclusive). Se dice que A se descompone en bloques cuando se la expresa en función de los $p \cdot q$ bloques A_{ij}; éstos forman p filas y q columnas y quedan situados como si se tratara de elementos de una matriz:

$$A = \left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \cdots & A_{1q} \\ \hline A_{21} & A_{22} & \cdots & A_{2q} \\ \hline \vdots & \vdots & \cdots & \vdots \\ \hline A_{p1} & A_{p2} & \cdots & A_{pq} \end{array}\right]$$

MULTIPLICACION POR BLOQUES

Sean dadas dos matrices $A_{m\times p}$ y $B_{p\times n}$ y considérese su producto $C_{m\times n}$. Supongamos que A se descompone en los bloques A_{ih} que determinan los índices de filas $(0 < m_1 < m_2 < \cdots < m_\alpha = m)$ y de columnas $(0 < p_1 < p_2 < \cdots < p_\gamma = p)$ y supongamos, también, que B está descompuesto en los bloques B_{hj} que determinan los índices de filas $(0 < p_1 < p_2 < \cdots < p_\gamma = p)$ y de columnas $(0 < n_1 < n_2 < \cdots < n_\beta = n)$. Pues bien, se verifica que el producto $C = AB$ puede expresarse en bloques como el resultado de multiplicar A por B operando con los bloques A_{ik} y B_{hj} como si se tratara de escalares. Más exactamente:

$$[A_{ih}] \cdot [B_{hj}] = [C_{ij}]$$
$$(A_{ih},\ B_{hj} \text{ y } C_{ij} \text{ bloques})$$

donde C_{ij} es el siguiente bloque (para $i = 1, 2, \ldots, \alpha$ y $j = 1, 2, \ldots, \beta$):

$$C_{ij} = \sum_{h=1}^{\gamma} A_{ih}B_{hj} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{i\gamma}B_{\gamma j}$$

COMPROBACION

Expresemos la matriz A como suma de las matrices A^*_{ir} que se obtiene de mantener, en A, el bloque A_{ir} y sustituir por ceros a todos los demás elementos. Análogamente, expresemos B como la suma de las correspondientes matrices B^*_{sj}. Haciendo esto, para el producto AB se obtiene:

$$\left.\begin{aligned} A &= \sum_{i,r} A^*_{ir} \\ B &= \sum_{s,j} B^*_{sj} \end{aligned}\right\} AB = \sum_{i,r,s,j} A^*_{ir}B^*_{sj} \qquad \begin{pmatrix} i = 1, 2, \ldots, \alpha \\ r,\ s = 1, 2, \ldots, \gamma \\ j = 1, 2, \ldots, \beta \end{pmatrix}$$

$$A^*_{ir}B^*_{sj} = \begin{bmatrix} O & \cdots & O & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & \cdots & A_{ir} & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & \cdots & O & \cdots & O \end{bmatrix} \begin{bmatrix} O & \cdots & O & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & \cdots & B_{sj} & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & \cdots & O & \cdots & O \end{bmatrix} = \begin{bmatrix} O & \cdots & O & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & \cdots & P_{ij} & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & \cdots & O & \cdots & O \end{bmatrix}$$

Veamos lo que se puede decir acerca del producto $A^*_{ir}B^*_{sj}$. Si es $r \neq s$, entonces este producto es la matriz nula pues al multiplicar cualquier fila de la 1.ª por cualquier columna de la segunda se obtiene una suma de sumandos todos nulos. Si es $r = s$, en el producto sólo pueden aparecer elementos no nulos en el bloque P_{ij}, de lugar (i, j), cuyos elementos son (salvo sumandos nulos) los productos de las filas de A_{ir} por las columnas de B_{rj}, o sea, ha de ser $P_{ij} = A_{ir}B_{rj}$. Al sumar todos estos productos $A^*_{ir}B^*_{sj}$, resulta que el bloque de lugar (i, j) es, pues, la suma e todos los posibles P_{ij}, que es el resultado que se había anunciado para C_{ij}.

β MATRICES INVERTIBLES

Nos vamos a ocupar ahora de una cuestión de gran utilidad: se trata del estudio de aquellas matrices que, siendo cuadradas, tienen inversa, así como de sus propiedades y de las de sus inversas.

3.6. DEFINICION Y PRIMERAS PROPIEDADES

[027]

Una matriz cuadrada A, de tamaño $n \times n$, se dice que es *invertible* (o inversible) si existe otra matriz, de igual tamaño, que se llama *matriz inversa* de A y se denotará por A^{-1}, tal que:

$$AA^{-1} = A^{-1}A = I \quad (I = \text{matriz unidad})$$

Si A y B son matrices cuadradas, de tamaño $n \times n$, se verifica:

1. Si A tiene inversa, entonces sólo tiene una inversa. Con mayor generalidad, también se verifica que, si dos matrices A' y A'', de tamaño $n \times n$, son tales que $AA' = I$, y $A''A = I$, entonces A tiene inversa y es $A^{-1} = A' = A''$.
2. Si A y B tienen inversa, entonces el producto AB también tiene inversa, que vale:
$$(AB)^{-1} = B^{-1}A^{-1}$$
3. Si A tiene inversa, entonces su traspuesta A^t también tiene inversa, que vale:
$$(A^t)^{-1} = (A^{-1})^t$$

COMPROBACIONES

1. Supongamos que A' y A'', de tamaño $n \times n$, son tales que $AA' = A''A = I$; entonces será:

$$A'' = A''I = A''(AA') = (A''A)A' = IA' = A'$$

Por tanto, $A' = A''$ y esta matriz satisface a lo exigido para ser matriz inversa de A. Por otra parte, nótese que si A' y A'' fuesen, ambas, inversas de A se verificaría que $AA' = I$ y $A''A = I$ y, según se acaba de ver, habría de ser $A' = A''$, luego todas las posibles inversas de A son iguales.

2. Aquí hay que comprobar que al multiplicar, por la derecha y por la izquierda, la matriz AB por la matriz $B^{-1}A^{-1}$, los dos productos que resultan son iguales a la matriz unidad. Así ocurre:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

3. Debemos comprobar que al multiplicar A^t por $(A^{-1})^t$ se obtiene la unidad I, tanto si se multiplica por la derecha como si se hace por la izquierda; esto es cierto, ya que:

$$A^t(A^{-1})^t = [A^{-1}A]^t = I^t = I$$

$$(A^{-1})^tA^t = [AA^{-1}]^t = I^t = I$$

EJERCICIO

Sean A y B dos matrices invertibles. En función de la matriz M^{-1}, inversa de $M = AB$, que se supone conocida, determínense A^{-1} y B^{-1}.

RESOLUCION

Las matrices invertibles de tamaño $n \times n$ forman grupo multiplicativo. Esto, que ocurre con los elementos invertibles de un anillo unitario cualquiera, significa: 1) que el producto de matrices invertibles es, a su vez, invertible; 2) que la matriz unidad es invertible; 3) que el producto es asociativo, y 4) que si A es invertible, entonces también lo es A^{-1}.

Sabemos que

$$M^{-1} = (AB)^{-1} = B^{-1}A^{-1}$$

Si premultiplicamos por B y, luego, posmultiplicamos por A, resulta que

$$BM^{-1} = B(B^{-1}A^{-1}) = (BB^{-1})A^{-1} = IA^{-1} = A^{-1}$$

$$M^{-1}A = (B^{-1}A^{-1})A = B^{-1}(A^{-1}A) = B^{-1}I = B^{-1}$$

esto es, las inversas pedidas son $A^{-1} = BM^{-1}$ y $B^{-1} = M^{-1}A$.

ALGUNAS PROPIEDADES RELATIVAS A LA INVERSA

A causa de su ulterior utilidad, reseñamos aquí algunas propiedades relacionadas con la matriz inversa; se trata de las siguientes:

[028]

1. Si A es una matriz cuadrada que tiene una fila o una columna nula, entonces A no es invertible.
2. Si $A_1, A_2, ..., A_p$ son matrices invertibles, del mismo tamaño, entonces también es invertible su producto $A_1A_2...A_p$ y se verifica que:

$$(A_1A_2...A_p)^{-1} = A_p^{-1}...A_2^{-1}A_1^{-1}$$

3. Al aplicar a la matriz unidad I (de cualquier tamaño) una operación elemental E, ya sea a sus filas ya a sus columnas, la matriz $E(I)$ que así se obtiene es invertible.

COMPROBACION

1. Si A tiene la fila i-ésima nula, entonces para cualquiera que sea la matriz B, de igual tamaño, la matriz AB también tiene nula su fila i-ésima, luego $AB \neq I$. Por ello, A no puede tener inversa. Análogamente, si la columna i-ésima de A es nula, entonces BA tiene nula su columna i-ésima, para cualquier matriz cuadrada B, de igual tamaño que A, luego $AB \neq I$ y, por ello, A no tiene inversa.
2. Aplicando reiteradamente la anterior propiedad 2 de [027], se puede poner:

$$(A_1A_2...A_p)^{-1} = [A_1 \cdot (A_2...A_p)]^{-1} = (A_2...A_p)^{-1}A_1^{-1} =$$
$$= [A_2(A_3...A_p)]^{-1} \cdot A_1^{-1} = (A_3...A_p)^{-1}A_2^{-1}A_1^{-1} = ... = A_p^{-1}...A_2^{-1}A_1^{-1}$$

3. Sólo consideraremos el caso de operaciones elementales en filas, esto es, supondremos que es $E = E^f$; para operaciones elementales en columnas se puede razonar de modo análogo.

Vamos a empezar suponiendo que E^f es una cualquiera de las tres operaciones elementales básicas, que en la demostración anterior (la de [026]) habíamos llamado $\dot{E}^f$, $\ddot{E}^f$ y $\dddot{E}^f$ y también $[i \leftrightarrow j]$, $[i \rightarrow \lambda i]$ con $\lambda \neq 0$ y $[i \rightarrow i + j]$; es decir, comenzaremos probando que las matrices elementales $\dot{E}^f(I)$, $\ddot{E}^f(I)$ y $\dddot{E}^f(I)$ son invertibles. Esto es evidente ya que estas tres matrices admiten como inversas respectivas a las matrices elementales que se obtienen al aplicar, a la unidad I, las operaciones elementales $[j \leftrightarrow i]$. $[i \rightarrow (1/\lambda)i]$ y $[i \rightarrow i - j]$, como se comprueba trivialmente.

En el caso general, la operación $E^f(I)$ es el resultado de aplicarle, a I, sucesivas operaciones elementales básicas $E_1^f, E_2^f, ..., E_h^f$ (de uno de los tres tipos anteriores), por lo que $E^f(I) = F_h...F_2F_1I$, donde $F_i = E_i^f(I)$ para $i = 1, 2, ..., h$.

Ahora bien, según se acaba de decir, todas estas matrices F_i son invertibles

y, según la anterior propiedad 2, de ello se desprende que también lo es su producto, esto es, que $E^f(I)$ es invertible, como deseábamos comprobar.

3.7. CALCULO EFECTIVO DE LA INVERSA

Realizando operaciones elementales en las filas o en las columnas de una matriz invertible, y si se procede de manera atinada, siempre se puede transformar la matriz en la unidad. Este resultado se utiliza para obtener un algoritmo que nos permita hallar, fácilmente, la matriz inversa.

[029]

1. Si M es la matriz escalonada (escalonada por columnas) que se obtiene aplicando unas ciertas operaciones elementales en filas (en columnas) a una matriz cuadrada A, entonces todos los elementos de la diagonal de M son no nulos sí y sólo si A es invertible.
2. Si A es una matriz cuadrada invertible, entonces su matriz canónica de equivalencia es la matriz unidad. A esta matriz se puede llegar, a partir de A, realizando operaciones elementales, en A, solamente a filas o, también, solamente a columnas. Si E^f y E^c son operaciones elementales que aplicadas, respectivamente, a las filas o a las columnas de A conducen a la unidad I, es decir, tales que $E^f(A) = I$ y $E^c(A) = I$, entonces:

$$A^{-1} = E^f(I) \quad \text{y} \quad A^{-1} = E^c(I)$$

3. La inversa A^{-1}, de una matriz invertible A, se puede hallar del siguiente modo: se realizan adecuadas operaciones elementales en A y, las mismas operaciones, en la matriz unidad I, y todo ello se hace sólo en filas o sólo en columnas, hasta conseguir que A se transforme en I, lo cual siempre es posible; en ese momento, la transformada de I es A^{-1}.

COMPROBACION

Vamos a considerar el caso de filas; para el de columnas se razona de manera análoga.

1. Llamemos E^f a la operación elemental que transforma A en M, esto es, tal que $M = E^f(A) = E^f(I) \cdot A$. Si A es invertible, como también lo es $E^f(I)$ (véase [028], 3), nos encontramos con que el producto de ambas, que es M, es invertible; por tanto, M no puede tener ninguna fila nula (véase [028], 1) y, en consecuencia, todos los elementos de su diagonal son no nulos (si alguno fuese nulo, su última fila sería nula). Recíprocamente, si los elementos de la diagonal de M son todos no nulos, se puede seguir el proceso

que se indica en el próximo apartado 2, que nos conduce a la existencia de A^{-1}, como había que comprobar.

2. Supongamos que, a partir de A, hemos encontrado ya la anterior matriz escalonada M, que tiene no nulos todos los elementos de su diagonal. Realizamos ahora nuevas operaciones elementales, en las filas de M, para llegar a la matriz unidad. Esto es fácil: basta con tomar como pivotes a los sucesivos elementos de la diagonal principal, empezando por el último y avanzando hacia el primero; cada vez que se utiliza uno de ellos como pivote, se le puede convertir, a él, en un 1 (si no lo es ya) y transformar en ceros a todos los elementos situados sobre él, al tiempo que se conservan los ceros ya conseguidos anteriormente. Es claro que así se llega a la matriz unidad.
3. Llamando E^f a la operación elemental resultante de todas las anteriores, nos encontramos con que $I = E^f(A) = E^f(I) \cdot A$. Haciendo otro tanto con las columnas y llamando E^c a la correspondiente operación elemental, se verificará que $I = E^c(A) = A \cdot E^c(I)$. Si denotamos $A' = E^f(I)$ y $A'' = E^c(I)$, los anteriores resultados se expresan poniendo $A'A = I$ y $AA'' = I$, luego de ello se desprende (véase [027], 1) que existe A^{-1} y que es $A^{-1} = A' = A''$, o sea que

$$E^f(I) = E^c(I) = A^{-1}$$

con lo que concluye esta comprobación.

EJEMPLO

Para hallar la matriz inversa de la siguiente matriz A, procediendo como se acaba de indicar, esto es, realizando operaciones elementales en filas, a partir de las matrices A e I o, lo que es igual, a partir de la matriz $[A \mid I]$, se obtiene:

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 4 & 0 \\ 3 & 5 & -1 \end{bmatrix} \; ; \quad [A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 & 1 & 0 \\ 3 & 5 & -1 & 0 & 0 & 1 \end{array}\right] \begin{matrix} \\ \rightarrow 1^a \\ \rightarrow 2^a \\ \rightarrow 3^a \end{matrix}$$

$$\begin{matrix} 1^{a'} = 1^a \rightarrow \\ 2^{a'} = 2^a - 2 \cdot 1^a \rightarrow \\ 3^{a'} = 3^a - 3 \cdot 1^a \rightarrow \end{matrix} \left[\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 4 & -2 & 1 & 0 \\ 0 & -4 & 5 & -3 & 0 & 1 \end{array}\right]$$

$$\begin{matrix} 1^{a''} = 1^{a'} \rightarrow \\ 2^{a''} = (-1/2)2^{a'} \rightarrow \\ 3^{a''} = 3^{a'} - 2 \cdot 2^a \rightarrow \end{matrix} \left[\begin{array}{ccc|ccc} 1 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & -1/2 & 0 \\ 0 & 0 & -3 & 1 & -2 & 1 \end{array}\right]$$

$$\begin{matrix} 1^{a'''} = 1^{a''} - (2/3) \cdot 3^{a''} \rightarrow \\ 2^{a'''} = 2^{a''} - (2/3) \cdot 3^{a''} \rightarrow \\ 3^{a'''} = (-1/3) \cdot 3^{a''} \rightarrow \end{matrix} \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & 1/3 & 4/3 & -2/3 \\ 0 & 1 & 0 & 1/3 & -5/6 & -2/3 \\ 0 & 0 & 1 & -1/3 & 2/3 & -1/3 \end{array}\right]$$

$$\begin{array}{ll} 1^{aIV} = 1^{a'''} - 3 \cdot 2^{a'''} \rightarrow & \\ 2^{aIV} = 2^{a'''} \rightarrow & \\ 3^{aIV} = 3^{a'''} \rightarrow & \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2/3 & 23/6 & 4/3 \\ 0 & 1 & 0 & 1/3 & -5/6 & -2/3 \\ 0 & 0 & 1 & -1/3 & 2/3 & -1/3 \end{array}\right] = [I \mid A^{-1}]$$

Como el primer bloque de esta última matriz es la unidad I, resulta que el segundo bloque es A^{-1}; por tanto:

$$A^{-1} = \frac{1}{6}\begin{bmatrix} -4 & 23 & 8 \\ 2 & -5 & -4 \\ -2 & 4 & -2 \end{bmatrix}$$

3.8. CARACTERIZACIONES DE LAS MATRICES INVERTIBLES

La existencia de la inversa de una matriz dada está íntimamente ligada a otras propiedades de ésta que tienen, también, gran interés. Nos referimos fundamentalmente a:

[030]

Una matriz cuadrada A, de tamaño $n \times n$, es invertible sí y sólo si se verifica una cualquiera de las siguientes condiciones (que resultan, por ello, equivalentes entre sí):

1. **Existe otra matriz B, de igual tamaño, tal que $AB = I$ y, en este supuesto, es $A^{-1} = B$.**
2. **Existe otra matriz B, de igual tamaño que A, tal que $BA = I$ y en este supuesto, es $A^{-1} = B$.**
3. **La matriz A tiene rango n.**
4. **La matriz A es *regular* o simplificable. Esto es, se verifica que: siempre que sea $AP = AQ$ o $PA = QA$, para algunas matrices P y Q de tamaño $n \times n$, entonces es $P = Q$. (Las matrices que no son regulares se llaman *singulares*.)**
5. **La matriz A no es *divisor de cero*. Se dice que una matriz $M \neq O$, de tamaño $n \times n$ es un divisor de cero si existe alguna matriz $N \neq O$, de tamaño $n \times n$, tal que $MN = O$ o $NM = O$.**

NOTA: **Más tarde veremos (en [040]) que las matrices regulares se caracterizan también por ser aquellas matrices cuadradas que tienen determinante no nulo.**

En un anillo cualquiera, se verifica que un elemento es regular si y sólo si no es divisor de cero. Si el anillo es unitario, entonces sus elementos invertibles son todos regulares; pero, en general, puede haber elementos regulares que no admiten inverso. Para el anillo unitario que forman las matrices $n \times n$ se verifica, además, que toda matriz regular es invertible.

DEMOSTRACIONES

1. Si A es invertible, entonces esta condición primera se cumple obviamente para $B = A^{-1}$. Si se cumple esta propiedad primera, entonces A ha de ser invertible, pues de no serlo se presentaría la siguiente contradicción: Re-

curramos a que siempre existe una operación elemental E^f tal que la matriz $M = E^f(A)$ es escalonada. Si A no fuese invertible, la matriz escalonada M tendría nulo alguno de los elementos de su diagonal (véase [029], 1), luego M tendría nula alguna de sus filas (la última al menos) y, como

$$MB = E^f(A)B = E^f(I)AB = E^f(I)I = E^f(I)$$

nos encontramos con que MB, que tiene alguna fila nula (por tenerla M), sería igual a $E^f(I)$, que es matriz invertible (véase [028], 3); como una matriz invertible no puede tener una fila nula, hemos dado con la contradicción anunciada.

2. Esta propiedad se prueba de análoga manera que la anterior.
3. Recurramos aquí, también, a que existe alguna operación elemental E^f tal que la matriz $M = E^f(A) = E^f(I)A$ es escalonada. Tengamos en cuenta que: M y A tienen el mismo rango (véase [019]). que M tiene rango n sí y sólo si todos los elementos de su diagonal son no nulos y que todos los elementos de la diagonal de M son no nulos sí y sólo si A es invertible (véase [029], 1). De estas tres propiedades se desprende que A tiene rango n sí y sólo si A es invertible.

4 y 5. Para comprobar estas dos propiedades, bastará con cerciorarse de que son ciertas las implicaciones *a*), *b*) y *c*) que figuran a continuación, o, lo que es igual, las *a*), *b*) y *c'*), ya que *c'*) equivale a *c*):

a) A invertible $\Rightarrow$ A regular.
b) A regular $\Rightarrow$ A no divisor de cero.
c) A no divisor de cero $\Rightarrow$ A invertible.
c') A no invertible $\Rightarrow$ A divisor de cero.

a) Si A es invertible, en el caso de ser $AP = AQ$ o $PA = QA$, multiplicando por A^{-1}, se obtiene:

$$AP = AQ \Rightarrow A^{-1}(AP) = A^{-1}(AQ) \Rightarrow IP = IQ \Rightarrow P = Q$$
$$PA = QA \Rightarrow (PA)A^{-1} = (QA)A^{-1} \Rightarrow PI = QI \Rightarrow P = Q$$

lo que prueba que A es regular.

b) Si A es regular y en el caso de ser $AP = O$ o $QA = O$, se tendrá:

$$AP = O \Rightarrow AP = AO \Rightarrow P = O$$
$$QA = O \Rightarrow QA = OA \Rightarrow Q = O$$

lo que prueba que A no es divisor de cero.

c') Empecemos recordando que, para cualquiera que sea A, existe una operación elemental E^f tal que la matriz $M = E^f(A)$ es escalonada. Si A no es invertible, entonces M tiene, al menos, su última fila nula (ya que algún elemento de la diagonal de M ha de ser nulo; véase [029], 1). Llamando P a la matriz que tiene todos sus elementos nulos excepto el de lugar (n, n) que vale 1, es evidente que $P \neq O$ y $PM = O$. Pues bien, si es $F = E^f(I)$,

como es $M = E^{f}(A) = FA$, premultiplicando a ambos miembros de esta igualdad por P, resulta que

$$O = PM = P(FA) = (PF)A$$

Pero PF es la matriz que tiene sus $n - 1$ primera filas nulas y su última fila es igual a la última fila de F, que no es nula, pues F es regular; por tanto $PF \neq O$ y $(PF)A = O$, lo que significa que PF es divisor de cero (por la izquierda). Nótese que cambiando, en el anterior razonamiento, filas por columnas, se obtiene que A es, también, divisor de cero por la derecha.

CAPITULO

4

Determinantes

Entre nuestros conocimientos, aunque no sean muchos y estén algo «oxidados», seguro que se encuentra aquello del volumen de un paralelepípedo. Recordemos que, si las aristas de éste vienen representadas por unos vectores $\vec{a} = (a_1, a_2, a_3)$, $\vec{b} = (b_1, b_2, b_3)$ y $\vec{c} = (c_1, c_2, c_3)$ (en ejes cartesianos rectangulares), entonces su volumen resulta ser el siguiente número V:

$$V = a_1b_1c_1 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1 - a_2b_1c_3 - a_1b_3c_2 =$$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

donde el último miembro es una forma abreviada de expresar el segundo, al que se ha llamado «determinante» de la matriz que tiene por columnas a los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$. Recuérdese también que si se cambia el orden de los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$, puede alterarse el signo de V, de ahí que a V se le llame algo así como «volumen orientado» o «volumen con signo». Este número V, no sólo nos da, pues, el volumen, que vale $|V|$ (valor absoluto de V), sino que también nos informa de la «orientación» que se ha dado al paralelepípedo, esto es, del orden en el que se toman sus aristas.

Pues bien, se va a generalizar el concepto de determinante, al caso de matrices cuadradas de cualquier tamaño $n \times n$. y se va a hacer de manera que sigan verificándose ahora las propiedades de las que goza el volumen orientado. Así, pues, se va a pedir que: 1) Si a una arista (la arista $\vec{a}$) se la prolonga o reduce multiplicando por una cantidad λ (la nueva arista es $\lambda\vec{a}$), entonces el volumen queda multiplicado por λ; 2) Si una arista (la arista $\vec{a}$) se descompone en suma de dos (se pone $\vec{a} = \vec{a}_1 + \vec{a}_2$), entonces el volumen resulta ser la suma de los volúmenes de los paralelepípedos que (manteniendo las demás aristas) resultan de sustituir la arista que se descompone por cada uno de sus componentes (las $\vec{a}_1$ y $\vec{a}_2$); 3) Si dos aristas coinciden (por ejemplo, $\vec{a} = \vec{b}$), entonces el volumen es nulo; 4) El paralelepípedo cuyas aristas son los vectores de la referencia cartesiana, tiene volumen unidad. También se verificarán, como consecuencia, otras propiedades, que se irán detallando en su momento, de entre las que destacamos una ya citada anteriormente: Si al definir el paralelepípedo, lo que se hace dando sus aristas, éstas se dan en distinto orden por que se permutan dos de ellas, entonces el volumen cambia de signo.

α DEFINICION Y VALOR DE UN DETERMINANTE

Aquí, al menos al principio, vamos a precisar de una notación específica con la que pueda apreciarse, de manera clara y explícita, que una matriz está formada por filas o por columnas y en la que se señale quiénes son éstas. Empecemos, pues, con ello:

NOTACION

[031] En lo que sigue, una matriz cuadrada $A = [a_{ij}]$, de tamaño $n \times n$ y con los escalares a_{ij} de un cuerpo cualquiera, se denotará también recurriendo a sus columnas o a sus filas poniendo:

$$A = [a_{ij}] = [{}^{c}\bar{a}_1, {}^{c}\bar{a}_2, \ldots, {}^{c}\bar{a}_n] = \begin{bmatrix} {}^{f}\bar{a}_1 \\ {}^{f}\bar{a}_2 \\ \vdots \\ {}^{f}\bar{a}_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \\ \uparrow & \uparrow & & \uparrow \\ {}^{c}\bar{a}_1 & {}^{c}\bar{a}_2 & & {}^{c}\bar{a}_n \end{bmatrix} \begin{matrix} \leftarrow {}^{f}\bar{a}_1 \\ \leftarrow {}^{f}\bar{a}_2 \\ \\ \\ \leftarrow {}^{f}\bar{a}_n \end{matrix}$$

donde:

$$\begin{cases} a_{ij} = \text{elemento de lugar } (i, j) \text{ de } A \\ {}^{c}\bar{a}_i = \text{columna } i\text{-ésima de } A \\ {}^{f}\bar{a}_i = \text{fila } i\text{-ésima de } A \end{cases}$$

Cuando se quiera destacar la columna i-ésima (igual se puede hacer con filas) y no tenga especial importancia quiénes son las restantes columnas, se pondrá:

$$A = [\cdots {}^{c}\bar{a}_i \cdots] \quad \text{en lugar de} \quad A = [{}^{c}\bar{a}_1 \ \cdots \ {}^{c}\bar{a}_i \ \cdots \ {}^{c}\bar{a}_n]$$

Para señalar que la i-ésima columna de A es $\bar{u}$ (igual se procedería con filas), se pondrá:

$$A = [\cdots (\bar{i}^{a} = \bar{u}) \cdots]$$

4

4.1. DETERMINANTE DE UNA MATRIZ CUADRADA

La siguiente definición responde a lo ya dicho en la anterior introducción. Este modo de presentar las determinantes tiene obvias ventajas, frente al de dar, sin más, la expresión que proporciona el valor del determinante, $\det A$, de una matriz A. Aquí buscaremos una expresión para $\det A$ y lo haremos de modo que el valor que de ello resulte verifique a los requerimientos que, previamente, se han fijado. Con esta vía, constructiva, queda justificada la ya conocida fórmula del determinante de una matriz; no obstante, tiene el inconveniente de que, a posteriori, hay que comprobar que dicha fórmula, no sólo satisface a lo exigido, sino que es la única que lo hace.

[032]

Se llama determinante a una aplicación que a cada matriz cuadrada $A = [a_{ij}]$ le asigna un escalar, que se llama determinante de la matriz A y se denota poniendo

$$\det A \quad , \quad \det[{}^c\bar{a}_1, {}^c\bar{a}_2, \ldots, {}^c\bar{a}_n]^{(*)} \quad \text{ó} \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}^{(**)}$$

de manera que se cumpla, para cualesquiera índices i y j, que:

1.ª $\det[{}^c\bar{a}_1 \cdots {}^c\bar{a}_i' + {}^c\bar{a}_i'' \cdots {}^c\bar{a}_n] = \det[{}^c\bar{a}_1 \cdots {}^c\bar{a}_i' \cdots {}^c\bar{a}_n] + \det[{}^c\bar{a}_1 \cdots {}^c\bar{a}_i'' \cdots {}^c\bar{a}_n]$.

2.ª $\det[{}^c\bar{a}_1 \cdots \lambda {}^c\bar{a}_i \cdots {}^c\bar{a}_n] = \lambda \det[{}^c\bar{a}_1 \cdots {}^c\bar{a}_i \cdots {}^c\bar{a}_n]$ (λ escalar)

3.ª $\det[{}^c\bar{a}_1 \cdots (i^{\text{ésima}} = \bar{u}) \cdots (j^{\text{ésima}} = \bar{u}) \cdots {}^c\bar{a}_n] = 0$ (o sea, ${}^c\bar{a}_i = {}^c\bar{a}_j \Rightarrow \det A = 0$)

4.ª $\det I = 1$ (I = matriz unidad).

Luego probaremos que el determinante de toda matriz existe y es único.

En la presente definición se pueden sustituir las columnas ${}^c\bar{a}_i$ de A por sus filas; más adelante se verá que de ambos modos se llega a un mismo resultado.

Se llama *determinante de orden n* al determinante de una matriz de tamaño $n \times n$.

(*) Por ${}^c\bar{a}_i$ se denota a la columna de lugar i de la matriz A.

(**) Al determinante de A se le denota también poniendo $|A|$.

La aplicación determinante (de una matriz $n \times n$),

$$\det : \mathcal{M}_{n\times n}(K) \to K$$
$$A \mapsto \det A$$

se dice que es una *forma multilineal alternada*. Se llama forma porque su espacio de llegada es el cuerpo K de escalares. Se dice que es multilineal para indicar que es lineal respecto de cada una de las columnas (condiciones 1.ª y 2.ª). Se dice que es alternada para señalar que se verifica la propiedad *a* de [033]).

Antes de comprobar que los determinantes así definidos existen y que cada matriz tiene un único determinante, analicemos algunas de sus propiedades, en el supuesto de su existencia y unicidad:

[033]

Para los determinantes, como consecuencia de la anterior definición, se verifican las siguientes propiedades:

a) $\det[{}^c\bar{a}_1 \cdots (i^a = \bar{u}) \cdots (j^a = \bar{v}) \cdots {}^c\bar{a}_n] = -\det[{}^c\bar{a}_1 \cdots (i^a = \bar{v}) \cdots (j^a = \bar{u}) \cdots {}^c\bar{a}_n]$.

b) $\det[{}^c\bar{a}_1 \cdots (i^a = \bar{o}) \cdots {}^c\bar{a}_n] = 0$.

c) $\det[{}^c\bar{a}_1 \cdots (i^a = \sum_{j \neq i} \lambda_j {}^c\bar{a}_j) \cdots {}^c\bar{a}_n] = 0$ (λ_j escalares cualesquiera)

d) $\det[{}^c\bar{a}_1 \cdots (i^a = {}^c\bar{a}_i + \sum_{j \neq i} \lambda_j {}^c\bar{a}_j) \cdots {}^c\bar{a}_n] = \det[{}^c\bar{a}_1 \cdots {}^c\bar{a}_i \cdots {}^c\bar{a}_n]$.

DEMOSTRACION

a) Sustituyendo las filas i-ésima y j-ésima, ambas, por $\bar{u} + \bar{v}$, se obtiene una matriz para cuyo determinante se tiene:

$$0 \overset{(1)}{=} \det[\cdots (i^a = \bar{u} + \bar{v}) \cdots (j^a = \bar{u} + \bar{v}) \cdots] \overset{(2)}{=} \det[\cdots (i^a = \bar{u}) \cdots (j^a = \bar{u}) \cdots] +$$
$$+ \det[\cdots (i^a = \bar{u}) \cdots (j^a = \bar{v}) \cdots] + \det[\cdots (i^a = \bar{v}) \cdots (j^a = \bar{u}) \cdots] +$$
$$+ \det[\cdots (i^a = \bar{v}) \cdots (j^a = \bar{v}) \cdots] \overset{(3)}{=} \det[\cdots (i^a = \bar{u}) \cdots (j^a = \bar{v}) \cdots] +$$
$$+ \det[\cdots (i^a = \bar{v}) \cdots (j^a = \bar{u}) \cdots]$$

(las igualdades (1), (2) y (3) son consecuencia de las hipótesis 3.ª, 1.ª y 3.ª de [032], respectivamente). El que este último miembro sea nulo no es otra cosa que la igualdad *a)*, que había que comprobar.

b) Tomando $\lambda = 0$ en la hipótesis 2.ª de [032] se obtiene que:

$$\det[\cdots (i^a = \bar{o}) \cdots] = \det[\cdots (i^a = 0\bar{o}) \cdots] = 0 \det[\cdots (i^a = \bar{o}) \cdots] = 0$$

c) Recurriendo a las hipótesis 1.ª, 2.ª y 3.ª de [032], se obtiene:

$$\det[\cdots (i^a = \sum_{j \neq i} \lambda_j {}^c\bar{a}_j) \cdots] = \sum_{j \neq i} \det[\cdots (i^a = \lambda_j {}^c\bar{a}_j) \cdots] =$$
$$= \sum_{j \neq i} \lambda_j \det[\cdots (i^a = {}^c\bar{a}_j) \cdots] = \sum_{j \neq i} \lambda_j 0 = 0$$

d) De acuerdo con la hipótesis 1.ª de [032] y con la anterior propiedad *c)*, se obtiene:

$$\det[\cdots (i^a = {}^c\bar{a}_i + \sum_{j \neq i} \lambda_j {}^c\bar{a}_j) \cdots] =$$
$$= \det[\cdots (i^a = a_i) \cdots] + \det[\cdots (i^a = \sum_{j \neq i} \lambda_j {}^c\bar{a}_j) \cdots] = \det[\cdots {}^c\bar{a}_i \cdots] + 0$$

RESUMEN

Lo dicho hasta ahora sobre las determinantes, en la definición [032] y en las propiedades ya demostradas en [032], se puede expresar, también, diciendo:

[034]

Para el determinante $\det A$, de una matriz cuadrada A, se verifica lo que sigue. Si una columna de A es combinación lineal de otras columnas (en particular, si es nula o si es igual a otra columna), entonces $\det A = 0$. Si a una columna de A se la multiplica por un escalar λ, entonces $\det A$ queda multiplicado por λ. Si una columna de A se descompone en dos sumandos, entonces $\det A$ es igual a la suma de los determinantes de las dos matrices que resultan de sustituir, en A, aquella columna por cada uno de los sumandos. Si a una columna de A se le suma una combinación lineal de otras columnas, con ello no se altera el valor del $\det A$. Si dos columnas de A se permutan, entre sí, entonces $\det A$ cambia de signo. El determinante de la matriz unidad vale 1.

Según se probará después, las anteriores propiedades se verifican también para las filas; esto es, donde arriba dice columna, puede ponerse fila y, con ello, todo lo allí dicho sigue siendo cierto.

EJERCICIOS

1. Compruébese que el siguiente determinante Δ es nulo:

$$\Delta = \begin{vmatrix} a_1 - b_1 & b_1 - c_1 & c_1 - a_1 \\ a_2 - b_2 & b_2 - c_2 & c_2 - a_2 \\ a_3 - b_3 & b_3 - c_3 & c_3 - a_3 \end{vmatrix}$$

Así ocurre ya que, si a su tercera columna la sumamos la primera y la segunda, con lo que el determinante no varía, resulta:

$$\Delta = \begin{vmatrix} a_1 - b_1 & b_1 - c_1 & 0 \\ a_2 - b_2 & b_2 - c_2 & 0 \\ a_3 - b_3 & b_3 - c_3 & 0 \end{vmatrix} = 0$$

2. Compruébese que el siguiente determinante Δ es un múltiplo entero del número 41; para comprobarlo será de gran utilidad tener en cuenta que 4141, 943, 2993 y 9717 son múltiplos de 41:

$$\Delta = \begin{vmatrix} 4 & 1 & 4 & 1 \\ 0 & 9 & 4 & 3 \\ 2 & 9 & 9 & 3 \\ 9 & 7 & 1 & 7 \end{vmatrix}$$

La propiedad es cierta ya que, sumando a la 4.ª columna la 3.ª multiplicada por 10 más la 2.ª por 100 más la 1.ª por 1000, con lo que no varía el valor del determinante, resulta:

$$\Delta = \begin{vmatrix} 4 & 1 & 4 & 4141 \\ 0 & 9 & 4 & 943 \\ 2 & 9 & 9 & 2993 \\ 9 & 7 & 1 & 9717 \end{vmatrix} = \begin{vmatrix} 4 & 1 & 4 & 41 \times \alpha \\ 0 & 9 & 4 & 41 \times \beta \\ 2 & 9 & 9 & 41 \times \gamma \\ 9 & 7 & 1 & 41 \times \delta \end{vmatrix} =$$

$$= 41 \begin{vmatrix} 4 & 1 & 4 & \alpha \\ 0 & 9 & 4 & \beta \\ 2 & 9 & 9 & \gamma \\ 9 & 7 & 1 & \delta \end{vmatrix} = 41 \times n = \text{múltiplo de } 41$$

(α, β, γ, δ y n son unos ciertos números naturales.)

4.2. EXPRESION DEL VALOR DE UN DETERMINANTE

A partir de la anterior definición, vamos a obtener una expresión que proporcione el valor del determinante de una matriz en función de los elementos de ésta. Pero, antes de entrar en ello, conviene que empecemos recordando algo de lo que ya sabemos acerca de las permutaciones o sustituciones y de las trasposiciones.

SUSTITUCIONES (RECORDATORIO)

[035] Sea dado un conjunto de N elementos, que podemos suponer que es el $N = \{1, 2, ..., n\}$. Para permutar estos elementos, esto es, para colocarlos en distinto orden, bastará con recurrir a una aplicación biyectiva $\sigma : N \rightarrow N$, que representaremos escribiendo debajo de cada $i \in N$ su imagen $\sigma(i)$, es decir, poniendo:

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \downarrow & \downarrow & \cdots & \downarrow \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

Estas biyecciones de N en N se llaman *permutaciones* o *sustituciones* de n elementos. Con una de ellas, $\sigma: N \rightarrowtail N$, los elementos de N, que inicialmente se les supone formando la sucesión natural $(1, 2, ..., n)$, se consideran ahora formando la nueva sucesión $(\sigma(1), \sigma(2), ..., \sigma(n))$. En total, hay $n!$ sustituciones de n elementos; al conjunto de todas ellas se le denota por $\mathcal{S}_n$. La composición de dos sustituciones es otra sustitución; la biyección inversa σ^{-1}, de una sustitución σ, es también una sustitución(*).

Se llaman *trasposiciones* a aquellas sustituciones en las que, salvo dos elementos de N, que vamos a llamar i y j, todos los demás permanecen fijos (esto es, coinciden con su imagen); los dos elementos que varían se transforman cada uno de ellos en el otro, esto es $i \rightarrowtail j$ y $j \rightarrowtail i$. Toda sustitución puede expresarse como composición de trasposiciones. Se verifica, además, que si $\sigma = \tau_p \circ \cdots \circ \tau_1$ y $\sigma = \tau'_q \circ \cdots \circ \tau'_1$ son dos descomposiciones, de una sustitución σ, en producto (composición) de trasposiciones. entonces p y q tienen la misma paridad (ambos son pares o ambos son impares). Se dice que una sustitución es par o impar según que se descomponga, respectivamente, en un número par o impar de trasposiciones. Se llama *signo* de una sustitución σ al valor $\varepsilon(\sigma)$ definido mediante:

$$\varepsilon(\sigma) = \begin{cases} +1, & \text{si } \sigma \text{ es par} \\ -1, & \text{si } \sigma \text{ es impar} \end{cases}$$

PROPIEDAD

Veamos, en primer lugar, un resultado que utilizaremos aquí y en otros lugares:

[036] Si A es una matriz cuadrada, de tamaño $n \times n$, cuyas columnas son ${}^c\bar{a}_1, {}^c\bar{a}_2, ..., {}^c\bar{a}_n$, y para cualquiera que sea la sustitución $\sigma \in \mathcal{S}_n$, se verifica:

$$\det[(1.^a = {}^c\bar{a}_{\sigma(1)}), (2.^a = {}^c\bar{a}_{\sigma(2)}) \cdots (n.^a = {}^c\bar{a}_{\sigma(n)})] =$$
$$= \varepsilon(\sigma) \det[{}^c\bar{a}_1, {}^c\bar{a}_2, ..., {}^c\bar{a}_n]$$

DEMOSTRACION

En efecto: si σ fuese, en particular, una trasposición, esta propiedad sería la [033], *a*), ya demostrada. Si σ no es una trasposición, la podemos descomponer en trasposiciones y aplicar, sucesivamente, la propiedad [033], *a*) a cada una de ellas; como cada vez que hagamos esto se produce un cambio de signo en el determinante, al final se obtiene un determinante que es el de partida afectado del signo + ó − según que sea par o impar el número de trasposiciones, esto es, multiplicado por $\varepsilon(\sigma)$, como había que comprobar.

(*) El conjunto $\mathcal{S}_n$, de las sustituciones de n elementos, es un grupo respecto de la composición de aplicaciones, que se llama grupo simétrico o de las sustituciones de n elementos.

VALOR DE UN DETERMINANTE

[037]

El determinante $\det A$, de una matriz $A = [a_{ij}]$, de tamaño $n \times n$, es igual a cualquiera de las dos expresiones siguientes:

$$(*)\begin{cases} \det A = \sum_{\sigma \in \mathcal{J}_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} & \text{[I]} \\ \det A = \sum_{\rho \in \mathcal{J}_n} \varepsilon(\rho) a_{\rho(1)1} a_{\rho(2)2} \cdots a_{\rho(n)n} & \text{[II]} \end{cases}$$

Debido a que las anteriores expresiones [I] y [II] son iguales se puede asegurar que:

1. Para cualquiera que sea la matriz cuadrada A, si A^t es su traspuesta, se verifica que:

$$\det A^t = \det A$$

2. Todas las propiedades de los determinantes de las matrices que se refieren a sus columnas son también válidas si se las refiere a las filas y recíprocamente.

(*) El determinante de A es igual a la suma de todos los $n!$ productos* del tipo $\pm a_{1i} a_{2j} \cdots a_{nl}$ o, también, del tipo $\pm a_{s1} a_{t2} \cdots a_{vn}$ tales que: 1) en cada producto hay un elemento, y sólo uno, de cada fila de A y, al tiempo, hay un, y sólo un, elemento de cada columna de A; y 2) el signo ($\pm$) es + ó − según que sea par o impar, respectivamente, la permutación $(h, k, ..., l)$ de $(1, 2, ..., n)$.

DEMOSTRACION

Procederemos por sucesivas etapas:

a) Si $\det A$ existe (cosa que, según luego veremos, es verdad), entonces es igual a la expresión [II]. En efecto, poniendo los vectores columna de A en función de los vectores columna canónicos ${}^c\bar{e}_1$, ${}^c\bar{e}_2$, ..., ${}^c\bar{e}_n$, es decir, recurriendo a que (para $i = 1, 2, ..., n$):

$${}^c\bar{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} = a_{1i} \underbrace{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{{}^c\bar{e}_1} + a_{2i} \underbrace{\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}}_{{}^c\bar{e}_2} + \cdots + a_{ni} \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}}_{{}^c\bar{e}_n} =$$

$$= a_{1i}\,{}^c\bar{e}_1 + a_{2i}\,{}^c\bar{e}_2 + \cdots + a_{ni}\,{}^c\bar{e}_n = \sum_{j=1}^{n} a_{ji}\,{}^c\bar{e}_j$$

resulta que, de acuerdo con propiedades ya conocidas, se verifica:

$$\det A = \det [{}^c\bar{a}_1, {}^c\bar{a}_2, \ldots, {}^c\bar{a}_n] = \det \left[\sum_{h=1}^{n} a_{h1}{}^c\bar{e}_h, \sum_{k=1}^{n} a_{k2}{}^c\bar{e}_k, \ldots, \sum_{l=1}^{n} a_{ln}{}^c\bar{e}_l \right] =$$

$$= \sum_{h,k,\ldots,l=1}^{n} \det [a_{h1}{}^c\bar{e}_h, a_{k2}{}^c\bar{e}_k, \ldots, a_{ln}{}^c\bar{e}_l] =$$

$$= \sum_{h,k,\ldots,l=1}^{n} a_{h1}a_{k2}\cdots a_{ln} \det [(1.^a = {}^c\bar{e}_h)(2.^a = {}^c\bar{e}_k)\cdots(n^a = {}^c\bar{e}_l)]$$

Siempre que dos de los índices $h, k, \ldots, l$ sean iguales, este último determinante será nulo (la matriz tiene dos columnas iguales); por ello, en esta última expresión, los únicos sumandos que no son nulos son aquéllos en los que $(h, k, \ldots, l)$ es una permutación de $(1, 2, \ldots, n)$, esto es, cuando existe alguna sustitución ρ, de n elementos o sea de $\mathcal{S}_n$, tal que $\rho(1) = h$, $\rho(2) = k, \ldots, \rho(l) = n$. Por todo ello, podemos continuar poniendo (con la ayuda de la propiedad *a*) recién demostrada):

$$\det A = \sum_{\rho \in \mathcal{S}_n} a_{\rho(1)1}a_{\rho(2)2}\cdots a_{\rho(n)n} \det [{}^c\bar{e}_{\rho(1)}, {}^c\bar{e}_{\rho(2)}, \ldots, {}^c\bar{e}_{\rho(n)}] =$$

$$\sum_{\rho \in \mathcal{S}_n} a_{\rho(1)1}a_{\rho(2)2}\cdots a_{\rho(n)n}\, \varepsilon(\rho) \det [{}^c\bar{e}_1, {}^c\bar{e}_2, \ldots, {}^c\bar{e}_n]$$

Como este último determinante es el determinante de la matriz unidad, que vale 1, hemos llegado ya a la expresión [II]. Es decir, si la definición de determinante (véase [032]) es consistente, o sea, si proporciona realmente algún valor para $\det A$, éste es el de la fórmula [II].

b) Las expresiones [I] y [II] son iguales. Para comprobar esto, recurriremos a que: 1) para cualquiera que sea la sustitución $\rho \in \mathcal{S}_n$, su recíproca (o inversa) es también una sustitución $\sigma = \rho^{-1} \in \mathcal{S}_n$; 2) que cuando ρ recorre todo $\mathcal{S}_n$, también $\sigma = \rho^{-1}$ recorre todo $\mathcal{S}_n$ y recíprocamente, y 3) que ρ y $\sigma = \rho^{-1}$ tienen el mismo signo. Téngase también en cuenta, que en el producto $a_{\rho(1)1}a_{\rho(2)2}\cdots a_{\rho(n)n}$ (que interviene en II), el factor genérico $a_{\rho(p)p}$ se puede poner, recurriendo al índice $q = \rho(p)$, en la forma $a_{q\rho^{-1}(q)}$. De todo ello, si en la expresión [II] se reordenan los factores del producto $a_{\rho(1)1}a_{\rho(2)2}\cdots a_{\rho(n)n}$ de manera que sean los primeros índices, en lugar de los segundos, los que queden en orden natural, se tiene:

$$[\mathrm{II}] = \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho) a_{\rho(1)1}a_{\rho(2)2}\cdots a_{\rho(n)n} = \text{(reordenando)}$$

$$= \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho) a_{1\rho^{-1}(1)}a_{2\rho^{-1}(2)}\cdots a_{n\rho^{-1}(n)} \overset{(1)}{=} \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} = [I]$$

(1) Debido a que $\varepsilon(\rho) = \varepsilon(\rho^{-1})$ y llamando $\sigma = \rho^{-1}$: nótese que si ρ recorre $\mathcal{S}_n$, también σ recorre $\mathcal{S}_n$.

c) Nótese que aún no se ha probado que los determinantes existan; ¿quién nos asegura que los requerimientos que se impusieron en su definición son compatibles? Hasta este momento lo que se ha hecho es suponer que existe algún valor para $\det A$, que satisface a lo exigido en la definición, y comprobar que, entonces, ese hipotético valor de $\det A$ debería ser el que proporciona la fórmula [I], o la [II]. Estamos, pues, obligados a verificar que el valor que proporciona la expresión [I], o la [II], satisface a las condiciones 1.ª, 2.ª, 3.ª y 4.ª (véase [032]) de la definición de determinante. Así ocurre, ya que para el valor que proporciona la expresión [II], se verifica que:

1. $$\det[{}^c\bar{a}_1 \cdots {}^c\bar{a}'_i + {}^c\bar{a}''_i \cdots {}^c\bar{a}_n] = \sum_{\rho \in \mathcal{P}_n} \varepsilon(\rho) a_{\rho(1)1} \cdots (a'_{\rho(i)i} + a''_{\rho(i)i}) \cdots a_{\rho(n)n} =$$

$$= \sum_{\rho \in \mathcal{P}_n} \varepsilon(\rho) a_{\rho(1)1} \cdots a'_{\rho(i)i} \cdots a_{\rho(n)n} + \sum_{\rho \in \mathcal{P}_n} \varepsilon(\rho) a_{\rho(1)1} \cdots a''_{\rho(i)i} \cdots a_{\rho(n)n} =$$

$$= \det[{}^c\bar{a}_1 \cdots {}^c\bar{a}'_i \cdots {}^c\bar{a}_n] + \det[{}^c\bar{a}_1 \cdots {}^c\bar{a}''_i \cdots {}^c\bar{a}_n]$$

2. $$\det[{}^c\bar{a}_1 \cdots \lambda {}^c\bar{a}_i \cdots {}^c\bar{a}_n] = \sum_{\rho \in \mathcal{P}_n} \varepsilon(\rho) a_{\rho(1)1} \cdots (\lambda a_{\rho(i)i}) \cdots a_{\rho(n)n} =$$

$$= \lambda \sum_{\rho \in \mathcal{P}_n} \varepsilon(\rho) a_{\rho(1)1} \cdots a_{\rho(i)i} \cdots a_{\rho(n)n} = \lambda \det[{}^c\bar{a}_1 \cdots {}^c\bar{a}_i \cdots {}^c\bar{a}_n]$$

3. Si las columnas i-ésima y j-ésima de una matriz $A = [a_{ij}]$ son iguales, esto es, si $a_{ri} = a_{rj}$ para todo r, vamos a probar que $\det A = 0$. En dicho supuesto, vamos a agrupar de dos en dos a los términos del desarrollo [*I*], de manera que junto a uno cualquiera de ellos, como el

$$\pm a_{h1} a_{k2} \cdots a_{pi} \cdots a_{qj} \cdots a_{ln} \quad [(hk \cdots p \cdots q \cdots l) \text{ es permutación de } (12 \cdots n)],$$

se considera el que resulta de intercambiar los índices p y q, esto es, el:

$$\mp a_{h1} a_{k2} \cdots a_{qi} \cdots a_{pj} \cdots a_{ln} \quad [(hk \cdots q \cdots p \cdots l) \text{ es permutación de } (12 \cdots n)]$$

Como $a_{pi} = a_{pj}$ y $a_{qi} = a_{qj}$, resulta que los dos términos anteriores son iguales, salvo el signo; éste es distinto ya que las permutaciones $(hk \cdots p \cdots q \cdots l)$ y $(hk \cdots q \cdots p \cdots l)$ tienen distinta paridad (pues de una se pasa a la otra mediante la trasposición $p \rightarrowtail q$, $q \rightarrowtail p$). Así, pues, la suma de estos dos términos es nula. Repitiendo este emparejamiento con todos los términos de la expresión [I], se obtiene que el valor de $\det A$ es cero.

4. Consideremos ahora que la matriz $A = [a_{ij}]$ es la unidad I, es decir, que $a_{ij} = 1$ si $i = j$ y que $a_{ij} = 0$ si $i \neq j$. En este caso, la expresión [I] conduce al siguiente resultado

$$\det A = \sum_{\rho \in \mathcal{P}_n} \varepsilon(\rho) a_{\rho(1)1} a_{\rho(2)2} \cdots a_{\rho(n)n} = +a_{11} a_{22} \cdots a_{nn} + 0 + 0 + \cdots + 0 =$$

$$= 1 \times 1 \times \cdots \times 1 = 1$$

d) Sólo nos queda por comprobar que $\det A^t = \det A$. Esto es cierto ya que, llamando $A = [a_{ij}]$ y $A^t = [a'_{ij}]$, donde $a'_{ij} = a_{ji}$, resulta:

$$\det A^t \overset{(1)}{=} \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho) a'_{\rho(1)1} a'_{\rho(2)2} \cdots a'_{\rho(n)n} =$$

$$= \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho) a_{1\rho(1)} a_{2\rho(2)} \cdots a_{n\rho(n)} \overset{(2)}{=} \det A$$

(Las igualdades (1) y (2) proceden de las expresiones [II] y [I] del determinante.)

EJEMPLOS

1. Los determinantes de las matrices de tamaños 2×2 y 3×3, es decir, los determinantes de órdenes 2 y 3, tienen los siguientes desarrollos.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

(las reglas mnemotécnicas que se utilizan para recordar este resultado se conocen con el nombre de regla de Sarrus).

2. Los desarrollos de los determinantes de órdenes mayores que 3 conducen a expresiones larguísimas, muy poco manejables; por ello, para obtener el valor de un determinante de orden mayor que 3, se siguen otros métodos, de los que hablamos en otros lugares de este capítulo. Para hacernos una idea del «tamaño» de estos desarrollos, obtengamos el más corto de ellos, el de orden 4:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11}a_{22}a_{33}a_{44} - a_{11}a_{22}a_{34}a_{43} + a_{11}a_{23}a_{34}a_{42} - a_{11}a_{23}a_{32}a_{44} + a_{11}a_{24}a_{32}a_{43} - a_{11}a_{24}a_{33}a_{42} + a_{12}a_{21}a_{34}a_{43} - a_{12}a_{21}a_{33}a_{44} + a_{12}a_{23}a_{31}a_{44} - a_{12}a_{23}a_{34}a_{41} + a_{12}a_{24}a_{33}a_{41} - a_{12}a_{24}a_{31}a_{43} + a_{13}a_{21}a_{32}a_{44} - a_{13}a_{21}a_{34}a_{42} + a_{13}a_{22}a_{34}a_{41} - a_{13}a_{22}a_{31}a_{44} + a_{13}a_{24}a_{31}a_{42} - a_{13}a_{24}a_{32}a_{41} + a_{14}a_{21}a_{33}a_{42} - a_{14}a_{21}a_{32}a_{43} + a_{14}a_{22}a_{31}a_{43} - a_{14}a_{22}a_{33}a_{41} + a_{14}a_{23}a_{32}a_{41} - a_{14}a_{23}a_{31}a_{42}$$

3. El determinante de una matriz triangular superior (o inferior) es igual al producto de los elementos de su diagonal; esto es:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}a_{33}\cdots a_{nn}$$

Así ocurre ya que los demás términos del desarrollo de este determinante, por tener un elemento (único) de cada fila y, también, de cada columna, son todos ellos nulos.

4. Si la primera columa (o fila) de una matriz tiene todos sus elementos nulos excepto el de lugar (1, 1), su determinante vale:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

Pues los términos no nulos del desarrollo del primer determinante son respectivamente iguales a los del desarrollo del segundo multiplicador por a_{11}.

EJERCICIOS

1. Compruébese que el siguiente determinante, llamado *determinante de Vandermonde*, toma el valor que se indica:

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a & b & c & \cdots & k \\ a^2 & b^2 & c^2 & \cdots & k^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a^{n-1} & b^{n-1} & c^{n-1} & \cdots & k^{n-1} \end{vmatrix} = (b-a)(c-a)\cdots(k-a)\cdot (c-b)\cdots(k-b)\cdot \cdots\cdots \cdots\cdots \cdots$$

RESOLUCION

Restando a cada fila la anterior multiplicada por a, el determinante queda en la forma:

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & b-a & c-a & \cdots & k-a \\ 0 & b(b-a) & c(c-a) & \cdots & k(k-a) \\ 0 & b^{n-2}(b-a) & c^{n-2}(c-a) & \cdots & k^{n-2}(k-a) \end{vmatrix} =$$

$$= (b-a)(c-a)\cdots(k-a) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ b & c & \cdots & k \\ \vdots & \vdots & \cdots & \vdots \\ b^{n-2} & c^{n-2} & \cdots & k^{n-2} \end{vmatrix}$$

Repitiendo con este último determinante, y con los que se vayan obteniendo (que también son de Vandermonde), el proceso anterior se llega al resultado del enunciado.

2. Hállese el valor del siguiente determinante de orden n:

$$\Delta_n = \begin{vmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 0 \end{vmatrix}$$

Si a cada una de las $n-1$ primeras filas se la resta la última fila y, en el determinante que así se obtiene, a la última fila se la suman todas las demás filas, se llega a:

$$\Delta_n = \begin{vmatrix} -1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & -1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & -1 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{vmatrix} = \begin{vmatrix} -1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & -1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & -1 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & n-1 \end{vmatrix} =$$

$$= (-1)^{n-1}(n-1)$$

4.3. CALCULO EFECTIVO DE UN DETERMINANTE

Para hallar un determinante de orden n mayor que 3 no es aconsejable recurrir al desarrollo de éste como suma de $n!$ términos, cada uno de los cuales es un producto de n elementos. Esta vía es, salvo en casos triviales, larga, larguísima, enormemente tediosa. Los desarrollos de un determinante que luego veremos (desarrollo por elementos de una fila o columna), tampoco suelen proporcionar métodos numéricos que resuelvan el problema de modo eficaz. Lo aconsejable para calcular determinantes es hacer uso de las propiedades fundamentales de los determinantes, con las que introducíamos el tema (véase [032] y [033]), y aplicarlas con método y acertadamente para ir transformando el determinante en otros que sean cada vez más fáciles de calcular, hasta llegar a uno que se halle trivialmente (por ejemplo, el determinante de una matriz triangular). Nótese que, según se acaba de indicar el método que aquí se propone, para hallar det A, recurre a realizar adecuadas operaciones elementales en las filas y en las columnas de la matriz A.

DESCRIPCION DEL METODO

[038]

Para hallar el determinante de una matriz cuadrada $A = [a_{ij}]$, de tamaño $n \times n$, hágase lo que sigue:

a) Si $a_{11} \neq 0$, no hay aquí, en *a*), nada que hacer y se pasará directamente al apartado *b*). Si $a_{11} = 0$ y $a_{i1} \neq 0$ para algún i, permútense las filas 1.ª e i-ésima, con lo que el determinante cambia de signo y tiene el elemento de lugar (1, 1) distinto de cero. Si todas las a_{i1} son nulas, el problema ya está resuelto pues es $\det A = 0$, y no hay nada más que hacer.

b) Se puede ya suponer que queremos hallar un determinante Δ_n cuyo elemento de lugar (1, 1) es no nulo. A los elementos de la primera columna los llamaremos (para simplificar notación) $\alpha_1, \alpha_2, ..., \alpha_n$, siendo $\alpha_1 \neq 0$. Restémosle a la fila i-ésima lo que resulta de multiplicar la 1.ª fila por α_i/α_1 y hagamos esto para $i = 2, 3, ..., n$; con ello, no se ha alterado el valor del determinante, que sigue valiendo Δ_n, y los elementos de la primera columna pasan a ser $\alpha_1, 0, 0, ..., 0$. El determinante Δ_n se puede, pues, poner en la forma $\Delta_n = \alpha_1 \Delta_{n-1}$, donde Δ_{n-1} es el determinante, de orden $n = 1$, que resulta de suprimir la primera fila y la primera columna de Δ_n.

c) Con el determinante Δ_{n-1} se repite el proceso anterior y, así, se le reduce a un determinante Δ_{n-2}, de orden $n - 2$. Reiterando este proceso, se llega a un determinante de 3.er o de 2.º orden, que se calcula trivialmente.

Este procedimiento admite múltiples variantes que, según los casos, pueden llegar a abreviar el proceso. En particular, puede ser preferible que, en lugar de hacer la operación elemental $(f \cdot i^{(a)}) \to (f \cdot i^{(a)}) - (\alpha_i/\alpha_1)(f \cdot 1.^{a})$ (el determinante quedaría ahora multiplicado por α_1), con lo que se evitarían las divisiones por α_1, que hacen más largo el desarrollo del método.

EJEMPLO

Siguiendo los pasos que se acaban de indicar, para hallar el siguiente determinante, se puede proceder como se señala:

$$\Delta = \begin{vmatrix} 0 & 3 & -2 & 1 & -4 \\ 4 & -2 & 0 & 2 & 6 \\ 1 & -2 & 1 & 3 & 5 \\ 3 & -3 & -1 & 5 & 5 \\ 5 & -1 & -1 & 6 & 3 \end{vmatrix} \begin{matrix} \leftarrow(1) \\ \leftarrow(2) \\ \leftarrow(3) \\ \leftarrow(4) \\ \leftarrow(5) \end{matrix}$$

$$\begin{matrix} (1)' = (3) \to \\ (2)' = (2):2 \to \\ (3)' = (1) \to \\ (4)' = (4) \to \\ (5)' = (5) \to \end{matrix} \quad \Delta = -2 \begin{vmatrix} 1 & -2 & 1 & 3 & 5 \\ 2 & -1 & 0 & 1 & 3 \\ 0 & 3 & -2 & 1 & -4 \\ 3 & -3 & -1 & 5 & 5 \\ 5 & -1 & -1 & 6 & 3 \end{vmatrix}$$

$$\begin{matrix} (1)'' = (1)' \to \\ (2)'' = (2) - 2(1)' \to \\ (3)'' = (3)' \to \\ (4)'' = (4)' - 3(1)' \to \\ (5)'' = (5)' - 5(1)' \to \end{matrix} \quad \Delta = -2 \begin{vmatrix} 1 & -2 & 1 & 3 & 5 \\ 0 & 3 & -2 & -5 & -7 \\ 0 & 3 & -2 & 1 & -4 \\ 0 & 3 & -4 & -4 & -10 \\ 0 & 9 & -6 & -9 & -22 \end{vmatrix} =$$

$$= -2 \begin{vmatrix} 3 & -2 & -5 & -7 \\ 3 & -2 & 1 & -4 \\ 3 & -4 & -4 & -10 \\ 9 & -6 & -9 & -22 \end{vmatrix}$$

$$\begin{matrix} (2)''' = (2)'' \to \\ (3)''' = (3)'' - (2)'' \to \\ (4)''' = (4)'' - (2)'' \to \\ (5)''' = (5)'' - 3(2)'' \to \end{matrix} \quad \Delta = -2 \begin{vmatrix} 3 & -2 & -5 & -7 \\ 0 & 0 & 6 & 3 \\ 0 & -2 & 1 & -3 \\ 0 & 0 & 6 & -1 \end{vmatrix} = -6 \begin{vmatrix} 0 & 6 & 3 \\ -2 & 1 & -3 \\ 0 & 6 & -1 \end{vmatrix} =$$

$$= -6(-36 - 12) = 288$$

β PROPIEDADES: DESARROLLO E INVERSA

4.4. DETERMINANTE DE UN PRODUCTO: CONSECUENCIA

Aquí, la cuestión central es comprobar que el determinante (como aplicación que a cada matriz A le asocia el escalar $\det A$) es lineal respecto del producto, esto es, que el determinante de un producto de matrices es igual al producto de los determinantes de aquéllas. Hay quienes esta propiedad la resumen diciendo que el determinante es una función multiplicativa.

DETERMINANTE DEL PRODUCTO DE DOS MATRICES

[039]

Para cualesquiera que sean las matrices cuadradas A y B, del mismo tamaño, se verifica que

$$\det(AB) = (\det A)(\det B)$$

DEMOSTRACION

Llamemos M a la matriz producto, $AB = M = [m_{ij}]$, cuyo elemento de lugar (i, j) es $m_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$ (las matrices A y B se suponen de tamaño $n \times n$). Entonces, la columna de lugar j de M (para $j = 1, 2, \ldots, n$), en función de las columnas de A, es:

$${}^c\bar{m}_j = {}^c\bar{a}_1 b_{1j} + {}^c\bar{a}_2 b_{2j} + \cdots + {}^c\bar{a}_n b_{nj} = \sum_{p=1}^{n} {}^c\bar{a}_p b_{pj}$$

Resulta entonces que:

$$\det(AB) = \det M = \det[{}^c\bar{m}_1, {}^c\bar{m}_2, \ldots, {}^c\bar{m}_n] =$$

$$= \det\left[\sum_{h=1}^{n} {}^c\bar{a}_h b_{h1}, \sum_{k=1}^{n} {}^c\bar{a}_k b_{k2}, \ldots, \sum_{l=1}^{n} {}^c\bar{a}_l b_{ln}\right] =$$

$$= \sum_{h,k,\ldots,l=1}^{n} \det[{}^c\bar{a}_h, {}^c\bar{a}_k, \ldots, {}^c\bar{a}_l] b_{h1} b_{k2} \cdots b_{ln}$$

Si en $(h, k, ..., l)$ hubiera algún índice repetido el último determinante sería nulo (pues tendría dos columnas iguales); podemos, pues, suponer que $(h, k, ..., l)$ es una permutación de (1, 2, ..., n), es decir, que $h = \sigma(1)$, $k = \sigma(2)$, ..., $l = \sigma(n)$ para cierta sustitución σ, que recorre $\mathcal{S}_n$. Por ello y recurriendo, también a la propiedad [036], resulta que:

$$\det(AB) = \sum_{\sigma \in \mathcal{S}_n} \det[{}^c\bar{a}_{\sigma(1)}, {}^c\bar{a}_{\sigma(2)}, ..., {}^c\bar{a}_{\sigma(n)}]b_{\sigma(1)1}b_{\sigma(2)2}\cdots b_{\sigma(n)n} =$$

$$= \det[{}^c\bar{a}_1, {}^c\bar{a}_2, ..., {}^c\bar{a}_n]\left(\sum_{\rho \in \mathcal{S}_n} \varepsilon(\sigma)b_{\sigma(1)1}b_{\sigma(2)2}\cdots b_{\sigma(n)n}\right) =$$

$$= \det A(\det B)$$

EJERCICIO

Para cualesquiera que sean las matrices cuadradas A y B, de igual tamaño, se verifica que

$$\det(AB) = \det(A^tB) = \det(AB^t) = \det(A^tB^t) =$$

$$= \det(BA) = \det(BA^t) = \det(BA^t) = \det(B^tA^t)$$

RESOLUCION

Como el determinante de una matriz es igual al de su traspuesta, todas las determinantes del primer renglón son iguales entre sí, por ser iguales a $(\det A)(\det B)$. Análogamente, los del segundo renglón son iguales a $(\det B)(\det A)$, que es igual al producto anterior.

DETERMINANTE NO NULO Y MATRIZ INVERSA

Las matrices regulares, de tamaño $n \times n$, han resultado ser (véase [030]) las que tienen inversa (son invertibles) y, también, aquellas cuyo rango es n. Pues bien, vamos a obtener aquí recurriendo a los determinantes, otra caracterización de las matrices regulares:

[040]

Una matriz cuadrada A es regular (o invertible) si y sólo si su determinante es distinto de cero; dicho de otro modo: A es singular (no regular) si y sólo si $\det A = 0$. Si A es regular, entonces

$$\det(A^{-1}) = \frac{1}{\det A}$$

DEMOSTRACION

Veamos primero que, si $\det A \neq 0$, entonces A es regular. Así es ya que, si A fuese singular, entonces el rango de A sería menor que n, luego A tendría una columna que sería combinación lineal de las demás, y, consecuentemente, sería $\det A = 0$, que va contra la hipótesis.

Veamos ahora que, si A es regular, entonces se verifica que $\det A \neq 0$. Como A es regular, existe su inversa A^{-1}, siendo $AA^{-1} = I$. Entonces, como el determinante de un producto es igual al producto de los determinantes, de $AA^{-1} = I$ se deduce que

$$(\det A)(\det A^{-1}) = \det I = 1 \qquad [1]$$

Por tanto, $\det A$ y $\det A^{-1}$ son no nulos (si alguno fuese nulo, lo sería su producto) y, en particular, es $\det A \neq 0$. Finalmente, despejando $\det A^{-1}$ de [1], se obtiene la igualdad del enunciado.

EJERCICIO

Analícese si hay alguna matriz cuadrada A de tamaño n impar que sea regular y tal que $A^t = -A^3$ (considerar los casos de A real y A compleja).

RESOLUCION

Tomando determinantes, de $A^t = -A^3$ se desprende que $\det A = (-1)^n (\det A)^3 = -(\det A)^3$. Como A debe ser regular, es $\det A \neq 0$, luego $(\det A)^2 = -1$. En el caso real no hay, pues, solución. En el caso complejo sí hay soluciones; por ejemplo:

$$A = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}$$

4.5. DESARROLLO POR LOS ELEMENTOS DE UNA LINEA

Seleccionemos una línea cualquiera (fila o columna) de una matriz cuadrada $A = [a_{ij}]$, de tamaño $n \times n$; supongamos que dicha línea es la fila i-ésima. En el desarrollo de su determinante, $\det A = \Sigma \pm a_{1k} a_{2k} \cdots a_{nl}$, dos elementos cualesquiera de la línea elegida no pueden figurar, ambos, en un mismo término $a_{1k} a_{2k} \cdots a_{nl}$ del desarrollo; en cada uno de estos términos aparece como factor uno, y sólo uno, de dichos elementos. Sacando cada uno de los referidos elementos a_{i1}, a_{i2}, ..., a_{in}, de la fila i-ésima, como factor común en los sumandos en los que figura, se llega a que

$$\det A = a_{i1} c_1 + a_{i2} c_2 + \cdots + a_{in} c_n$$

para unos ciertos coeficientes $c_1, c_2, ..., c_n$. Pues bien vamos a ocuparnos ahora en buscar quiénes son estos coeficientes y, consecuentemente, a obtener una nueva expresión para $\det A$.

DESARROLLO DE UN DETERMINANTE POR UNA LINEA

[041]

Sea $A = [a_{ij}]$ una matriz cuadrada, de tamaño $n \times n$. Para cada elemento a_{ij} de A, se conviene en llamar:

- *menor* de a_{ij} es el escalar $m_{ij} = \det A_{ij}$, donde A_{ij} es la submatriz que resulta de suprimir, en A, la fila i-ésima y la columna j-ésima.
- *Adjunto* o *cofactor* de a_{ij} es el escalar $\alpha_{ij} = (-1)^{i+j} \det A_{ij}$

Para cualquiera que sea el índice $i (i = 1, 2, ..., n)$, se verifica que:

$$\det A = a_{i1}\alpha_{i1} + a_{i2}\alpha_{i2} + \cdots + a_{in}\alpha_{in} \left(\begin{array}{l} \text{desarrollo por elementos} \\ \text{de la fila } i\text{-ésima de A} \end{array} \right)$$

$$\det A = a_{1i}\alpha_{1i} + a_{2i}\alpha_{2i} + \cdots + a_{ni}\alpha_{ni} \left(\begin{array}{l} \text{desarrollo por elementos} \\ \text{de la columna } i\text{-ésima} \end{array} \right)$$

DEMOSTRACION

Probaremos la validez del desarrollo por elementos de una columna (la i-ésima); análogamente se puede proceder para el caso de una fila. Para ello, descompongamos dicha columna como sigue:

$${}^c\bar{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} = a_{1i} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_{2i} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + a_{ni} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = a_{1i}{}^c\bar{e}_1 + a_{2i}{}^c\bar{e}_2 + \cdots + a_{ni}{}^c\bar{e}_n$$

con lo que se obtiene que $\det A$ se puede poner:

$$\det A = \det [{}^c\bar{a}_1 \cdots {}^c\bar{a}_i \cdots {}^c\bar{a}_n] = \det \left[{}^c\bar{a}_1 \cdots \sum_{h=1}^{n} a_{hi}\, {}^c\bar{e}_h \cdots {}^c\bar{a}_n \right] =$$

$$= \sum_{h=1}^{n} a_{hi} \det [{}^c\bar{a}_1 \cdots (i^{a} = {}^c\bar{e}_h) \cdots {}^c\bar{a}_n] = a_{1i}c_1 + a_{2i}c_2 + \cdots + a_{ni}c_n$$

donde

$$c_h = \det\left[{}^c\bar{a}_1 \cdots (i^a = {}^c\bar{e}_h) \cdots {}^c\bar{a}_n\right] = \begin{vmatrix} a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{h1} & a_{h2} & \cdots & 1 & \cdots & a_{hn} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = \det A'_{hi}$$

(columna i^a)

(la matriz A'_{hi} se obtiene sustituyendo, en A, la columna i-ésima por una columna de $n-1$ ceros y un uno en el lugar h-ésimo).

En consecuencia, sólo hay que comprobar que $c_h = \det A'_{hi}$ es igual al adjunto α_{hi} del enunciado (para $h = 1, 2, ..., n$) esto es, que

$$\det A'_{hi} = (-1)^{h+i} \det A_{hi} \quad (A_{hi} = \text{submatriz adjunta del elemento } a_{hi}, \text{ en } A)$$

Vamos a ello. En la matriz A'_{hi} intercambiemos la columna i-ésima por la columna que le precede y repitamos, sucesivamente, esta operación hasta llevar a la columna i-ésima al primer lugar. de modo que las restantes columnas quedan en el mismo orden relativo que tenían en A'_{hi}, y en A. Nótese que se ha realizado un total de $i-1$ cambios de columnas; por ello, llamando A''_{hi} a la matriz que así se obtiene, se verifica que $\det A'_{hi} = (-1)^i \det A''_{hi}$. En la matriz A''_{hi} intercambiemos ahora la fila h-ésima por la fila que tiene delante y repitamos, sucesivamente, esta operación hasta llevar a la fila h-ésima al primer lugar, de modo que las restantes filas quedan en el mismo orden relativo que tenían en A''_{hi}, y en A'_{hi}, y en A. Nótese que se ha realizado un total de $h-1$ cambios de filas; por ello, llamando A'''_{hi} a la matriz que así se obtiene, se verifica que $\det A''_{hi} = (-1)^h \det A'''_{hi}$. Conviene destacar que si en A'''_{hi} se prescinde de la 1.ª fila y de la 1.ª columna, la submatriz que se obtiene no es otra que la A_{hi}, que se cita en el enunciado. De todo lo dicho, resulta que:

$$\det A'_{hi} = (-1)^i \det A''_{hi} = (-1)^{i+h} \det A'''_{hi} = (-1)^{i+h} \det \left[\begin{array}{c|ccc} 1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & A_{hi} & \\ 0 & & & \end{array}\right] =$$

$$= (-1)^{i+h} \det A_{hi} = \alpha_{hi}$$

(los $*$ denotan a elementos de la fila h-ésima de A)

con lo que concluye la demostración.

EJEMPLO

Para el siguiente determinante, al desarrollarse por los elementos de su segunda columna, se obtiene:

$$\begin{vmatrix} 1 & 3 & 2 & 1 \\ 2 & 1 & -1 & 4 \\ 3 & 2 & -1 & 6 \\ 5 & -4 & 2 & -1 \end{vmatrix} = 3(-1)^{1+2}\begin{vmatrix} 2 & -1 & 4 \\ 3 & -1 & 6 \\ 5 & 2 & -1 \end{vmatrix} + 1(-1)^{2+2}\begin{vmatrix} 1 & 2 & 1 \\ 3 & -1 & 6 \\ 5 & 2 & -1 \end{vmatrix} +$$

$$+ 2(-1)^{3+2}\begin{vmatrix} 1 & 2 & 1 \\ 2 & -1 & 4 \\ 5 & 2 & -1 \end{vmatrix} + (-4)(-1)^{4+2}\begin{vmatrix} 1 & 2 & 1 \\ 2 & -1 & 4 \\ 3 & -1 & 6 \end{vmatrix} =$$

$$= -3(2 + 24 - 30 + 20 - 3 - 24) + 1(1 + 6 + 60 + 5 + 6 - 12) -$$

$$-2(1 + 4 + 40 + 5 + 4 - 8) - 4(-6 - 2 + 24 + 3 + 4 - 24) =$$

$$= 33 + 66 - 92 + 4 = 11$$

CONSECUENCIAS

A partir del desarrollo de una matriz por los elementos de una línea, se deducen las siguientes propiedades:

[042]

Si $A = [a_{ij}]$ es una matriz cuadrada, de tamaño $n \times n$, y llamando α_{ij} al adjunto o cofactor de su elemento a_{ij} (para $i, j = 1, 2, ..., n$), se verifica que:

(I) La suma de los productos de los elementos de una línea de A (fila o columna) por los adjuntos de los respectivos elementos de una línea paralela a aquella, es cero; es decir:

$$i \neq j \;\Rightarrow\; \begin{cases} a_{i1}\alpha_{j1} + a_{i2}\alpha_{j2} + \cdots + a_{in}\alpha_{jn} = 0 & \text{(para las filas)} \\ a_{1i}\alpha_{1j} + a_{2i}\alpha_{2j} + \cdots + a_{ni}\alpha_{nj} = 0 & \text{(para columnas)} \end{cases}$$

(II) Llamando δ_{ij} a los *símbolos de Kronecker* o elementos de la matriz unidad (esto es, $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ si $i \neq j$), es:

$$\left.\begin{aligned} a_{i1}\alpha_{j1} + a_{i2}\alpha_{j2} + \cdots + a_{in}\alpha_{jn} &= (\det A)\delta_{ij} \\ a_{1i}\alpha_{ij} + a_{2i}\alpha_{2j} + \cdots + a_{ni}\alpha_{nj} &= (\det A)\delta_{ij} \end{aligned}\right\} = \begin{cases} \det A & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

DEMOSTRACION

(I) Consideremos el caso de filas; para columnas se razonaría de igual manera. Recurramos a la matriz auxiliar $B = [b_{ij}]$ que coincide con la A salvo en la fila j; la fila j de B es igual a la fila i de A, de manera que B tiene iguales sus filas i y j. Nótese que, en B, los adjuntos $\beta_{j1}, \beta_{j2}, \ldots, \beta_{jn}$, de los elementos de su fila j, coinciden con los respectivos adjuntos $\alpha_{j1}, \alpha_{j2}, \ldots, \alpha_{jn}$ de los elementos de la fila j de A. Como B tiene dos filas iguales, su determinante es nulo; desarrollando éste por los elementos de su fila j, se obtiene que

$$0 = \det B = b_{j1}\beta_{j1} + b_{j2}\beta_{j2} + \cdots + b_{jn}\beta_{jn} =$$
$$= a_{i1}\alpha_{j1} + a_{i2}\alpha_{j2} + \cdots + a_{in}\alpha_{jn}$$

(II) Para $i = j$, el primer miembro es el desarrollo de $\det A$ por los elementos de la línea (fila o columna) i, luego la igualdad es cierta. Para $i \neq j$, se acaba de probar que el primer miembro es 0, luego la igualdad también es cierta en este caso.

4.6. MATRIZ INVERSA

A partir de los últimos resultados obtenidos en el apartado anterior, que se han resumido en [042], se puede obtener fácilmente la expresión que da la matriz inversa A^{-1} de una matriz regular dada A.

EXPRESION DE LA MATRIZ INVERSA

[043]

Sea $A = [a_{ij}]$ una matriz cuadrada, de tamaño $n \times n$. Se llama *matriz adjunta* (o de cofactores) de A a la matriz $^*A = [\alpha_{ij}]$, de tamaño $n \times n$, cuyos elementos son las α_{ij} = (adjunto de a_{ij}). Se verifica que:

(I) $A \cdot {}^*A^t = {}^*A^t \cdot A = (\det A)I$ (I = unidad $n \times n$)

(II) $(\det A)(\det {}^*A) = (\det A)^n$

(III) Si A es regular, es decir, si $\Delta = \det A \neq 0$, su inversa A^{-1} vale:

$$A^{-1} = \left(\frac{1}{\Delta}\right) {}^*A^t = \begin{vmatrix} \frac{\alpha_{11}}{\Delta} & \frac{\alpha_{21}}{\Delta} & \cdots & \frac{\alpha_{n1}}{\Delta} \\ \frac{\alpha_{12}}{\Delta} & \frac{\alpha_{22}}{\Delta} & \cdots & \frac{\alpha_{n2}}{\Delta} \\ \vdots & \vdots & & \vdots \\ \frac{\alpha_{1n}}{\Delta} & \frac{\alpha_{2n}}{\Delta} & \cdots & \frac{\alpha_{nn}}{\Delta} \end{vmatrix}$$

DEMOSTRACION

Llamemos $*a_{ij}$ al elemento de lugar (i, j) de la matriz $*A^t$; esto es, se pone $*a_{ij} = \alpha_{ji}$.

(I) De acuerdo con la definición de producto de matrices y recurriendo a la anterior «consecuencia» (II) de [042], se concluye que los elementos del producto $A \cdot *A^t = [c_{ij}]$ son:

$$c_{ij} = \sum_{h=1}^{n} a_{ih} *a_{hj} = \sum_{h=1}^{n} a_{ih}\alpha_{jh} = a_{i1}\alpha_{j1} + a_{i2}\alpha_{j2} + \cdots + a_{in}\alpha_{jn} = (\det A)\,\delta_{ij}$$

luego, como δ_{ij} es el elemento de lugar (i, j) de la matriz unidad I, resulta que $A \cdot *A^t = (\det A)I$, como había que comprobar. Análogamente se procede con el otro producto, el $*A^t \cdot A$.

(II) Tomando determinantes en la igualdad $A \cdot *A^t = (\det A)I$, ya probada en (I), se obtiene que:

$$(\det A) \cdot (\det *A^t) = \det[(\det A)I] = (\det A)^n \det I = (\det A)^n$$

como $\det *A^t = \det *A$, queda así probada la igualdad (II).

(III) Ahora que $\det A \neq 0$, la igualdad (I) se puede poner, dividiendo por $\Delta = \det A$, en la forma:

$$A \cdot \left(\frac{*A^t}{\Delta}\right) = \left(\frac{*A^t}{\Delta}\right) \cdot A = I$$

lo que significa que la matriz $*A^t/\Delta$ es la inversa de A, como había que comprobar.

EJERCICIO

Sea A una matriz regular que tiene todos sus elementos enteros. Pruébese que los elementos de A^{-1} son también números enteros si y sólo si $\det A = \pm 1$.

RESOLUCION

Como los elementos de A son enteros, también lo son sus adjuntos (ya que son sumas de productos de elementos de A), esto es, la matriz adjunta $*A$ tiene todos sus elementos enteros. Llamemos $\Delta = \det A$.

- Si $\Delta = \pm 1$, como $A^{-1} = *A^t/\Delta = \pm *A^t$ y $*A$ tiene todos sus elementos enteros, resulta evidente que también los tiene A^{-1}.

• Supongamos ahora que A^{-1} tiene todos sus elementos enteros. En este supuesto, será:

$$\left.\begin{array}{l} A^{-1} = {}^*A^t/\Delta \\ (\det A)(\det {}^*A) = (\det A)^n \end{array}\right\} \Rightarrow \left.\begin{array}{l} {}^*A^t = \Delta A^{-1} \\ \Delta(\det {}^*A) = \Delta^n \end{array}\right\} \Rightarrow$$

$$\Rightarrow \Delta[\det(\Delta A^{-1})] = \Delta^n \Rightarrow \Delta[\Delta^n \det A^{-1}] = \Delta^n \Rightarrow \det A^{-1} = \frac{1}{\Delta}$$

Como A^{-1} tiene todos sus elementos enteros, entonces $\det A^{-1}$ es entero, es decir, $1/\Delta$ es entero, lo que obliga a que sea $\Delta = \pm 1$.

OBSERVACION

Para calcular, de modo efectivo, la inversa de una matriz regular, no es aconsejable recurrir a la anterior expresión de la inversa ($A^{-1} = {}^*A^t/\Delta$), pues ello conduce a procedimientos excesivamente laboriosos. La inversa se calcula con mucha mayor eficacia utilizando el método ya descrito en [029].

4.7. RANGO DE MENORES DE UNA MATRIZ

Cuando, tiempo atrás (véase [018]), se definió el rango de una matriz, se empezó hablando del rango de sus filas y del rango de sus columnas y, luego, comprobamos que eran iguales. Consideramos ahora otro concepto de rango, el rango de menores de una matriz, y comprobaremos que viene a coincidir con el rango ya conocido; puede, pues, suponerse que el rango por menores no es otra cosa que una nueva forma de presentar el mismo rango que hemos venido manejando hasta ahora.

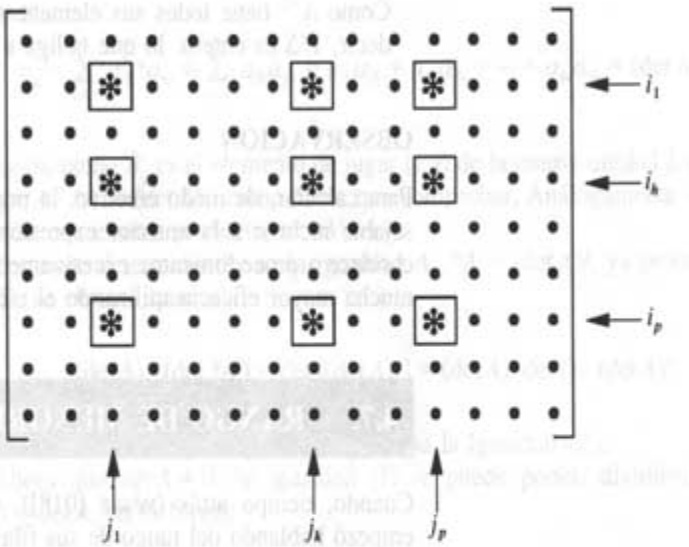
MENORES DE UNA MATRIZ

[044]

Dada una matriz $A = [a_{ij}]$, de tamaño $m \times n$ cualquiera, y elegidas las p filas $i_1, i_2, \ldots, i_p$ y las p columnas $j_1, j_2, \ldots, j_p$ de A, (con $p \leq m$ y $p \leq n$), se llama *menor*, de orden p de A, que determinan las p filas y las p columnas elegidas, al determinante de la submatriz de A, de tamaño $p \times p$, que forman los elementos $a_{i_k j_k}$, situados en los cruces de las filas y columnas elegidas; esto es, al determinante:

$$M = \begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_p} \\ \cdots\cdots & & \cdots\cdots \\ \cdots\cdots & & \cdots\cdots \\ a_{i_p j_1} & \cdots & a_{i_p j_p} \end{vmatrix}$$

Conviene insistir en que los elementos que intervienen en el menor M son (sólo) aquellos que están ubicados en alguna de las p líneas elegidas y, también, se hallan en alguna de las p columnas elegidas; esto es, se trata de los elementos situados en las p^2 intersecciones de las citadas filas y columnas, que son los que se señalan con asteriscos encerrados en las cuadrículas del siguiente esquema:



RANGO DE MENORES

Dada una matriz A, de tamaño $m \times n$, consideremos todos sus menores: los de orden 1, los de orden 2, ... (hasta llegar al menor de los m y n). De todos ellos, nos van a interesar los que son no nulos, que serán de distintos órdenes. Pues bien, al mayor de estos órdenes (de menores no nulos) se le llamará rango de menores de la matriz A. Es decir, este rango es p si hay alguna submatriz de tamaño $p \times p$, de A, que sea regular y todas las submatrices cuadradas de mayor tamaño que p son singulares. Se verifica que este rango no es, en esencia, algo nuevo sino, más bien, otro modo de definir el rango:

[045]

Se dice que p es el «rango de menores» de una matriz A, de tamaño $m \times n$ cualquiera, si A tiene algún menor de orden p que no es nulo y todos los menores de A de orden mayor que p son nulos; o sea, p es el mayor de los órdenes de los menores no nulos de A.

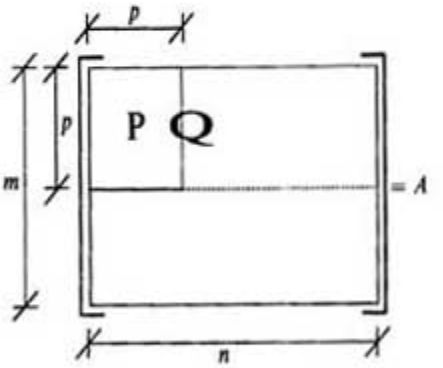
El «rango de menores» de una matriz A, cualquiera, es igual al rango de la matriz $r = \text{rang}\, A$.

DEMOSTRACION

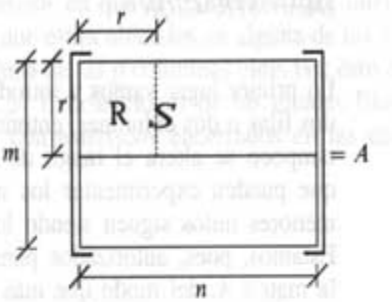
En primer lugar vamos a introducir una simplificación. Si se cambian entre sí dos filas o dos columnas, entonces $r = \text{rang}\, A$ no varía, según ya sabemos, pero tampoco se altera el rango de menores, p, puesto que la única modificación que pueden experimentar los menores es un cambio en su signo, luego los menores nulos siguen siendo los mismos y, en consecuencia, p no se altera. Estamos, pues, autorizados para permutar filas y para permutar columnas, en la matriz A, del modo que más nos cuadre.

Comprobaremos primero que es $p \leqslant r$ y después que es $r \leqslant p$.

- [Se verifica que $p \leqslant r$.] Sabemos que hay un menor de orden p de A que es no nulo; según lo que se ha dicho al principio, podemos suponer que dicho menor es el determinante de la submatriz P formada por las p primeras filas y las p primeras columnas de A. Como P tiene determinante no nulo, sus p columnas son linealmente independientes. Recurramos ahora a la submatriz Q que forman las p primeras filas de A (el tamaño de Q es $p \times n$). Las p primeras columnas de Q, por ser las columnas de p, son linealmente independientes, luego el rango de Q es al menos p y, por ello, sus p filas son linealmente independientes. Como estas p filas son filas de A, nos encontramos con que A tiene, al menos, p filas independientes, luego el rango de A es al menos p, es decir, se verifica que $p \leqslant r$.



- [Se verifica que $r \leqslant p$.] Sabemos que en A hay r filas linealmente independientes; podemos suponer, de acuerdo con lo que se dijo al comienzo, que dichas filas son las r primeras. Llamemos S a la submatriz que forman estas r primeras filas (el tamaño de S es $r \times n$); el rango de S es r. Por esto, sabemos que en S hay r columnas linealmente independientes; también podemos suponer ahora que dichas columnas son las r primeras. Recurramos a la submatriz R que forman las r primeras filas y las r primeras columnas de A; esta matriz R, de tamaño $r \times r$, tiene sus r columnas linealmente independientes y, en consecuencia, tiene determinante no nulo. Por tanto $\det R$ es un menor de orden r no nulo; como p es el mayor de los órdenes de los menores no nulos, se ha de verificar que $r \leqslant p$.



BUSQUEDA DEL RANGO SIRVIENDOSE DE LOS MENORES

Debemos empezar diciendo que el procedimiento eficaz para hallar el rango de una matriz numérica es el ya descrito en [019], recurriendo a las operaciones elementales. No obstante, también se puede hallar el rango recurriendo a los menores: buscando uno que sea no nulo y tal que todas las de mayor orden sean nulos. Este camino, que a primera vista parece ya muy largo, puede reducirse de modo apreciable si se procede metódicamente.

Es obvio que si se descubre, de un vistazo, que dos líneas paralelas son proporcionales, se puede prescindir de una de ellas sin que, con ello, se varíe el rango. Salvo que todas las líneas sean proporcionales, no cuesta nada encontrar un menor de orden 2 que sea no nulo.

[046]

Para hallar el rango de A, se toma un menor M_2 de orden 2 no nulo y se le orla con una fila fija, la i, y con sucesivas columnas; si todos los menores de orden 3 que así se obtienen son nulos, entonces se prescinde de la fila i, pues es combinación lineal de las de M_2, y se repite el proceso con otra o con otras filas hasta : 1) encontrar un menor M_3 de orden 3 no nulo, en cuyo caso el rango es al menos 3; ó 2) descubrir que todos los menores de orden 3 son nulos, en cuyo caso el rango es 2. Si hay un menor M_3 no nulo, se le orla con una fila y con sucesivas columnas, siguiendo el mismo proceso que con M_2, lo que nos lleva o bien a que el rango es 3 (si todos los menores de orden 4 son nulos) o bien a que el rango es al menos 4 (en cuanto se encuentre un menor de orden 4 no nulo). Siguiendo así, se llega a un menor no nulo del mayor tamaño posible; este tamaño es el rango buscado.

γ TEOREMAS DE CRAMER Y DE ROUCHE

Las cuestiones de álgebra lineal las iniciábamos nosotros hablando de los sistemas de ecuaciones lineales; ello nos ha sido bastante provechoso, nos ha permitido avanzar con buen pie. El método de Gauss, que allí fue el plato

fuerte, ha dado muchos más frutos de los que pudiera haberse esperado. No obstante, se nos quedaron allí bastantes cosas en el tintero, pues no disponíamos aún de excesivos conocimientos y, además, no era entonces momento oportuno para profundizar en estos asuntos.

Ha llegado el momento de reanudar el estudio de los sistemas de ecuaciones lineales; echando mano de lo que ya sabemos acerca de los vectores, las matrices, el rango y los determinantes, podemos obtener importantes resultados teóricos, que ya nos están haciendo falta.

NOTACION

Vamos a hacer aquí algunas puntualizaciones sobre la forma de denotar a los sistemas de ecuaciones lineales. Se va a presentar a éstos de modo que se pueda apreciar fácilmente su relación con los vectores y con las matrices:

[047]

Un sistema de ecuaciones lineales, como el $[S]$, se expresará también de cualquiera de las maneras $[S']$ ó $[S'']$, donde $A = [a_{ij}]$ es la matriz de coeficientes, $\mathcal{A}\mathcal{B} = [a_{ij} \mid b_i]$ es la matriz del sistema, $X = [x_j]$ es la matriz columna de las incógnitas, ${}^c\bar{a}_1$, ${}^c\bar{a}_2$, ..., ${}^c\bar{a}_n$ son los vectores columna de A y ${}^c\bar{b}$ es el vector columna de los términos independientes:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\} \quad [S] \quad \left(\begin{array}{c}\text{Sistema de } m \text{ ecuaciones} \\ \text{lineales, con las } n \text{ incógnitas} \\ x_1,\ x_2,\ \ldots,\ x_n\end{array}\right)$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad [S'] \quad \left(\begin{array}{c}\text{Abreviadamente, se pone} \\ AX = B\end{array}\right)$$

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad [S''] \quad \left(\begin{array}{c}\text{Abreviadamente,} \\ \text{se pone} \\ {}^c\bar{a}_1x_1 + {}^c\bar{a}_2x_2 + \\ + \cdots + {}^c\bar{a}_nx_n = {}^c\bar{b}\end{array}\right)$$

4.8. SISTEMAS DE CRAMER

Para empezar, analizaremos un caso fácil, útil y frecuente; se trata de un tipo de sistema de ecuaciones que es compatible y con solución única. Para él, vamos a obtener una fórmula que proporciona la solución. Hay que advertir desde el principio que el resultado de este teorema tiene mucha mayor utilidad teórica que práctica.

[048]

Un sistema de ecuaciones lineales $AX = B$ se dice que es un *sistema de Cramer* si su matriz de coeficientes A es cuadrada y regular: esto es, si el sistema tiene tantas ecuaciones como incógnitas y su determinante es no nulo (se llama determinante de este sistema a $\Delta = \det A$). Los sistemas de Cramer son, pues, los siguientes:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}\right\} \quad \text{con } \Delta \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0$$

Regla de Cramer. Todo sistema de Cramer tiene una y sólo una solución; esta es (para el anterior sistema $AX = B$):

$$x_1 = \frac{\Delta_1}{\Delta}\ ,\quad x_2 = \frac{\Delta_2}{\Delta}\ ,\quad \cdots\ ,\quad x_n = \frac{\Delta_n}{\Delta}\ ,$$

donde $\Delta_1, \Delta_2, \ldots, \Delta_n$ son los determinantes de las matrices que se obtienen al sustituir, en la matriz A del sistema, la columna 1.ª, 2.ª, ..., n.ª, respectivamente, por la columna de términos independientes.

DEMOSTRACION

El sistema, con notación matricial, es el $AX = B$, donde A es una matriz cuadrada regular, esto es, con inversa A^{-1}. El sistema es equivalente a $X = A^{-1}B$, ya que: 1) premultiplicando en el sistema por A^{-1}, se obtiene $X = A^{-1}B$; y 2) premultiplicando aquí por A, se obtiene el sistema. Podemos, pues, afirmar que el sistema tiene solución única y que ésta es $X = A^{-1}B$.

Recurriendo a la expresión de A^{-1} (véase [043]), la solución se puede poner:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = X = A^{-1}B = \frac{1}{\Delta}\begin{bmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{n2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} =$$

$$= \frac{1}{\Delta}\begin{bmatrix} \alpha_{11}b_1 + \alpha_{21}b_2 + \cdots + \alpha_{n1}b_n \\ \alpha_{12}b_1 + \alpha_{22}b_2 + \cdots + \alpha_{n2}b_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1n}b_1 + \alpha_{2n}b_2 + \cdots + \alpha_{nn}b_n \end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_n \end{bmatrix} \text{ (*)}$$

Por tanto, $x_i = \Delta_i/\Delta$ (para $i = 1, 2, ..., n$), como había que comprobar.

EJEMPLO

Aunque la Regla de Cramer no es un método práctico para hallar la solución (numérica) de un sistema (es preferible, salvo en casos triviales, el método de Gauss), a título de ejemplo vamos a resolver un sistema con la ayuda de la Regla de Cramer. Aplicando esta regla al siguiente sistema, se obtiene:

$$\left.\begin{aligned} 2x + y - 3z &= 1 \\ -x + 5y + z &= 4 \\ 3x - 2y - 4z &= -1 \end{aligned}\right\} \quad \Delta = \begin{vmatrix} 2 & 1 & -3 \\ -1 & 5 & 1 \\ 3 & -2 & -4 \end{vmatrix} = -4$$

$$x = \frac{\begin{vmatrix} 1 & 1 & -3 \\ 4 & 5 & 1 \\ -1 & -2 & -4 \end{vmatrix}}{\Delta} = 3 \; ; \; y = \frac{\begin{vmatrix} 2 & 1 & -3 \\ -1 & 4 & 1 \\ 3 & -1 & -4 \end{vmatrix}}{\Delta} = 1 \; ; \; z = \frac{\begin{vmatrix} 2 & 1 & 1 \\ -1 & 5 & 4 \\ 3 & -2 & -1 \end{vmatrix}}{\Delta} = 2$$

4.9. TEOREMA DE ROUCHE (**)

Una vez analizados los sistemas de Cramer, los más sencillos, vamos a apoyarnos en lo allí obtenido para estudiar un caso general cualquiera. Estamos interesados, aquí, en el estudio de la compatibilidad, y no en la búsqueda de las soluciones.

(*) El desarrollo de Δ_i (para $i = 1, 2, ..., n$) por los elementos de su columna i-ésima es $\Delta_i = \alpha_{1i}b_1 + \alpha_{2i}b_2 + \cdots + \alpha_{ni}b_n$.

(**) Este teorema admite una formulación más completa que la que ahora se considera. Más adelante se amplía lo que aquí se dice (véase [106]).

[049]

Considérese un sistema de ecuaciones lineales cualesquiera, como el siguiente:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\} \quad AX = B$$

Discusión del sistema. Dependiendo del número n de incógnitas y de los rangos de las matrices A y $\mathbb{AB}$ (de coeficientes y ampliada), se tiene:

1) $AX = B$ es compatible (con alguna solución) $\Leftrightarrow$ rang $\mathbb{AB}$ = rang A.
2) $AX = B$ tiene una y sólo una solución $\Leftrightarrow$ rang $\mathbb{AB}$ = rang $A = n$.
3) $AX = B$ tiene más de una solución $\Leftrightarrow$ rang $\mathbb{AB}$ = rang $A < n$.

DEMOSTRACION

Vamos a recurrir a los vectores columna de A y a la columna de B, que denotaremos por ${}^c\bar{a}_1$, ${}^c\bar{a}_2$, ..., ${}^c\bar{a}_n$ y ${}^c\bar{b}$, con lo que el sistema dado queda en la forma

$${}^c\bar{a}_1x_1 + {}^c\bar{a}_2x_2 + \cdots + {}^c\bar{a}_nx_n = {}^c\bar{b}$$

A la vista de esto, es evidente que el sistema tiene alguna solución si y sólo si ${}^c\bar{b}$ es una combinación lineal de ${}^c\bar{a}_1$, ${}^c\bar{a}_2$, ..., ${}^c\bar{a}_n$, lo que equivale a decir que la última columna de $\mathbb{AB}$ (esto es, la de B) es combinación lineal de las n primeras columnas de $\mathbb{AB}$ (esto es, de las columnas de A). Esto último, a su vez, significa que el rango de A no varía si se le añade la columna B, es decir, que rang A = rang $\mathbb{AB}$. Con esto queda probado el apartado 1).

Como los rangos de A y de $\mathbb{AB}$ o valen n o son menores que n, para comprobar los otros dos apartados bastará con ver que:

a) Si hay dos soluciones distintas, entonces es rang A = rang $\mathbb{AB} < n$.
b) Si rang A = rang $\mathbb{AB} < n$, entonces hay al menos dos soluciones.

a) Si hubiera dos soluciones distintas $x_i = \alpha_i$ y $x_i = \beta_i$ (para $i = 1, 2, ..., n$), se habría de verificar que:

$$\sum_{i=1}^{n} {}^c\bar{a}_i\alpha_i = {}^c\bar{b} \quad \text{y} \quad \sum_{i=1}^{n} {}^c\bar{a}_i\beta_i = {}^c\bar{b} \quad \text{luego (restando):}$$

$${}^c\bar{a}_1(\alpha_1 - \beta_1) + {}^c\bar{a}_2(\alpha_2 - \beta_2) + \cdots + {}^c\bar{a}_n(\alpha_n - \beta_n) = \bar{o}$$

Ahora bien, como alguno de los $\alpha_i - \beta_i$ ha de ser no nulo (pues las soluciones eran distintas), esta relación significa que las columnas de A son linealmente dependientes, es decir, que rang $A < n$. Además, por haber solución, es rang $\mathbb{AB}$ = rang A.

b) Por ser rang A = rang $\overline{A}$, hay alguna solución $x_i = \alpha_i$ (para $i = 1, 2, ..., n$); pues bien, por ser rang $A < n$, vamos a poder encontrar otra solución. Como el rango de A es menor que n (número de columnas de A), hay alguna combinación lineal de ellas (con coeficientes no todos nulos) que es igual a $\bar{o}$; esto es:

$$^c\bar{a}_1\lambda_1 + {}^c\bar{a}_2\lambda_2 + \cdots + {}^c\bar{a}_n\lambda_n = \bar{o} \quad \text{(para alguna } \lambda_i \text{ no todos nulos)}$$

Resulta entonces que

$$\sum_{i=1}^{n} {}^c\bar{a}_i(\alpha_i + \lambda_i) = \sum_{i=1}^{n} {}^c\bar{a}_i\alpha_i + \sum_{i=1}^{n} {}^c\bar{a}_i\lambda_i = {}^c\bar{b} + \bar{o} = {}^c\bar{b}$$

esto es, $x_i = \alpha_i + \lambda_i$ (para $i = 1, 2, ..., n$) es solución del sistema y es solución distinta de la $x_i = \alpha_i$, pues alguno de los λ_i es no nulo.

SOLUCIONES DE UN SISTEMA COMPATIBLE

[050]

Se considera aquí un sistema de ecuaciones lineales $AX = B$, que se supone compatible (rang A = rang $\overline{A}$):

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\}$$

Llamaremos:
r = rang A = rang $\overline{A}$

Se puede suponer(*) que el menor que forman las r primeras filas y las r primeras columnas de A es no nulo (menor principal); de ser así, el sistema $AX = B$ es equivalente al siguiente (llamado sistema principal):

$$\left.\begin{array}{l} a_{11}x_1 + \cdots + a_{1r}x_r = b_1 - (a_{1r+1}x_{r+1} + \cdots + a_{1n}x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1 + \cdots + a_{rr}x_r = b_r - (a_{rr+1}x_{r+1} + \cdots + a_{rn}x_n) \end{array}\right\}$$ Sistema principal

Para cualquiera que sean los valores que se les den a las incógnitas x_{r+1}, ..., x_n el sistema principal (que se convierte entonces en un sistema de Cramer) se satisface para unos únicos valores de las incógnitas x_1, ..., x_r (incógnitas principales), que dependen, pues, de los valores que se asignarán a las x_{r+1}, ..., x_n.

(*) Como r = rang A, hay algún menor de orden r de A que no es nulo. Si este menor no fuese el del enunciado, se podría conseguir que lo fuese sin más que cambiar el orden de colocación de las incógnitas y alterando los subíndices con los que se numera a las variables.

COMPROBACION

Las $m-r$ últimas filas de $\mathcal{AB}$ son combinaciones lineales de las r primeras, luego las $m-r$ últimas ecuaciones son combinaciones lineales de las r primeras; por ello, al prescindir de las $m-r$ últimas ecuaciones se obtiene un sistema (el principal) que es equivalente al $AX=B$.

Al dar valores arbitrarios a $x_{r+1}, ..., x_n$, el sistema principal se convierte en uno de r ecuaciones con r incógnitas cuyo determinante es no nulo (pues es el menor principal), esto es, pasa a ser un sistema de Cramer. Este sistema (para cada conjunto de valores que se tomen para $x_{r+1}, ..., x_n$) conduce, pues, a unos únicos $x_1, ..., x_r$.

CASO DE LOS SISTEMAS HOMOGENEOS

Al aplicar los resultados anteriores a los sistemas homogéneos, esto es, a los que tienen nulos sus términos independientes, se obtiene:

[051]

Considérese cualquier sistema de ecuaciones lineales que sea *homogéneo*, es decir, que tenga nulos todos sus términos independientes:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array}\right\} AX = O$$

(I) El sistema $AX=O$ es siempre compatible, pues tiene, al menos, la solución $x_1=0, x_2=0, ..., x_n=0$, que se llama *solución nula*.

(II) El sistema $AX=O$ tiene alguna solución no nula si, y sólo si, rang $\Delta <$ n.

(III) Si $AX=O$ tiene tantas ecuaciones como incógnitas (o sea, si $m=n$), entonces tendrá alguna solución no nula si, y sólo si, $\det A = 0$.

COMPROBACION

Respecto de un sistema general $AX=B$, la novedad estriba ahora en que es $B=O$ y, por ello, la relación rang $\mathcal{AB}$ = rang A se verifica siempre, luego $AX=O$ es compatible. Según sabemos, del teorema de Rouche, el sistema tendrá más de una solución si, y sólo si, es rang $\mathcal{AB}$ = rang $A < n$; como ahora la igualdad rang $\mathcal{AB}$ = rang A se cumple siempre, habrá más de una solución y, por ello, alguna solución no nula si rang $A < n$. Finalmente, si $m=n$, la condición rang $A < n$ equivale a la $\det A = 0$.

Ejercicios y problemas a la parte I

ENUNCIADOS

I.1. Resolver los sistemas de ecuaciones:

$$\left.\begin{aligned} x-2y+z&=5\\ 2x-y-2z&=-1\\ x+3y+z&=0 \end{aligned}\right\} \qquad \left.\begin{aligned} x+2y+2z&=0\\ 3x+y+3z&=3\\ 2x+y-2z&=-1\\ 2x+y+z&=2 \end{aligned}\right\}$$

I.2. Resolver, en el caso de ser compatibles, los sistemas de ecuaciones:

$$\left.\begin{aligned} 3x-4y+6z&=7\\ 5x+2y-4z&=5\\ x+3y-5z&=3 \end{aligned}\right\} \qquad \left.\begin{aligned} x+y-z&=3\\ 2x-y+4z&=3\\ 3x+2y-z&=8 \end{aligned}\right\}$$

Hallar los valores del parámetro α que hacen compatibles a los siguientes sistemas. Para dichos valores de α, resolver estos:

$$\left.\begin{aligned} x+3y+z&=0\\ 2x+y-3z&=5\\ -x+7y+9z&=\alpha \end{aligned}\right\} \qquad \left.\begin{aligned} x+y+z&=1\\ 2x-y+3z&=5\\ 3x+4y-2z&=2\\ x+3y-2z&=\alpha \end{aligned}\right\}$$

I.4. Resolver los siguientes sistemas de ecuaciones:

$$\left.\begin{aligned} x_1-x_2+2x_3-x_4&=0\\ 2x_1+x_2-x_3+x_4&=4\\ -x_1+2x_2+x_3-2x_4&=-2\\ 3x_1-x_2-3x_3+x_4&=1 \end{aligned}\right\}$$

b)
$$\left.\begin{aligned} x_1+x_2+x_3-x_4&=0\\ 2x_1+3x_2-x_3-3x_4&=0\\ -3x_1-2x_2-2x_3+4x_4&=0\\ 4x_1+2x_2+5x_3-2x_4&=1 \end{aligned}\right\}$$

Estudiar la compatibilidad del siguiente sistema, en función del parámetro α, y resolverlo cuando sea compatible:

$$\left.\begin{aligned} x+2y+4z-3u&=2\\ 3x+7y+5z-5u&=3\\ 5x+12y+6z-7u&=\alpha \end{aligned}\right\}$$

Resolver el siguiente sistema de ecuaciones:

$$\left.\begin{aligned} x+y+z-3u+2v&=4\\ x+2y-2z-5u+2v&=-3\\ 3x-y-3z-u+4v&=4 \end{aligned}\right\}$$

Resolver el siguiente sistema de ecuaciones:

$$\left.\begin{aligned} x-y+z-u&=8\\ 2x+4y+2z-u&=5\\ 3x+2y-z&=5\\ 2x+y-u&=5 \end{aligned}\right\}$$

Discutir, en función del parámetro α, el siguiente sistema de ecuaciones y resolverlo cuando sea compatible:

$$\left.\begin{aligned} x+y-z&=1\\ 3x+\alpha y+\alpha z&=5\\ 4x+\alpha y&=5 \end{aligned}\right\}$$

I.9. Sea S un sistema de m ecuaciones lineales con n incógnitas. Sabiendo que es $m<n$, pruébese que S puede ser incompatible o compatible indeterminado, pero no puede ocurrir que tenga una sola solución.

I.10. Resolver el sistema:

$$\left.\begin{aligned} x_1 + x_2 + x_3 - 2x_4 - 4x_5 &= -3 \\ x_1 + x_2 - x_3 + 2x_4 + 2x_5 &= 1 \\ x_1 + x_2 - x_3 - 2x_4 - 2x_5 &= -3 \\ x_1 + 5x_2 - 4x_3 + 3x_5 &= -1 \\ x_1 + 2x_2 - 2x_3 + 2x_4 + 3x_5 &= 1 \end{aligned}\right\}$$

I.11. Discutir (en función de α) y resolver cuando sea compatible el siguiente sistema de ecuaciones:

$$\left.\begin{aligned} \alpha x + y + z &= 1 \\ x + \alpha y + z &= \alpha \\ x + y + \alpha z &= \alpha^2 \end{aligned}\right\}$$

I.12. Si ninguno de los vectores $\vec{u}_1, \vec{u}_2, ..., \vec{u}_p$ depende linealmente del sistema $S = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_q\}$, ¿puede ocurrir que alguna combinación lineal de aquéllos dependa linealmente de S?

I.13. Estudiar si los siguientes sistemas de vectores de $\mathbb{R}^3$ son linealmente dependientes o independientes:

a) $\{(5, 3, 1), (1, 3, 2), (1, 1, 1)\}$

b) $\{(3, 2, 1), (1, -3, 2), (-1, -2, 3)\}$

I.14. Si $\bar{u}$ y $\bar{v}$ son dos vectores distintos, analizar si algunos de los vectores $(2 + \lambda)\bar{u} + (3 - \lambda)\bar{v}$ (λ = escalar) pueden ser iguales.

I.15. Sean $S = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_p\}$ y $T = \{\bar{u}_1, \bar{u}_2, ..., \bar{u}_q\}$ dos sistemas de vectores. Si cada uno de los S y T es linealmente independiente y todos los vectores de T son independientes de S, ¿el sistema $S \cup T$ ha de ser independiente?

I.16. Sea $S = \{\bar{u}_1, \bar{u}_2, ..., \bar{u}_p\}$ un sistema de vectores de $\mathbb{R}^n$. Si es $p > n$, compruébese que el sistema S es linealmente dependiente.

I.17. Estudiar si el sistema $S = \{\bar{u}, \bar{v}, \bar{w}\}$ de vectores de $\mathbb{C}^3$ es linealmente dependiente o independiente, en los siguientes casos:

a) $\bar{u} = (1, 0, -i)$, $\bar{v} = (0, i, -1)$, $\bar{w} = (i, 1, 1 + i)$

b) $\bar{u} = (1, i, -i)$, $\bar{v} = (0, 1, 1 + 2i)$, $\bar{w} = (1, 1 + i, -1)$

I.18. Sean dados los vectores

$$\bar{u} = (2, 3, 5) \quad , \quad \bar{v} = (1, 2, 3) \quad \text{y} \quad \bar{w} = (2, 1, 3)$$

Analizar si existen algunos escalares a_i, b_i y c_i tales que los vectores:

$$\bar{x} = a_1\bar{u} + b_1\bar{v} + c_1\bar{w}$$

$$\bar{y} = a_2\bar{u} + b_2\bar{u} + c_2\bar{w}$$

$$\bar{z} = a_3\bar{u} + b_3\bar{u} + c_3\bar{w}$$

sean linealmente independientes.

I.19. Sean $\bar{u}_1, \bar{u}_2, ..., \bar{u}_n$ los vectores de $\mathbb{R}^n$ que tienen los siguientes componentes: el vector $\bar{u}_i$ (para $i = 1, 2, ..., n$) tiene todas sus componentes valiendo 1 excepto la de lugar i que vale h. Hallar el rango de $\{\bar{u}_1, \bar{u}_2, ..., \bar{u}_n\}$ en función de h.

I.20. Hallar el rango del sistema:

$$S = \{\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4, \bar{u}_5, \bar{u}_6\}$$

siendo:

$$\bar{u}_1 = (2, 6, -1, 0, -1)$$

$$\bar{u}_2 = (2, -2, 1, -3, 2)$$

$$\bar{u}_3 = (3, 4, 1, -1, 2)$$

$$\bar{u}_4 = (1, -2, 2, -1, 2)$$

$$\bar{u}_5 = (-1, 0, 1, 2, 1)$$

$$\bar{u}_6 = (1, 8, -3, 1, -4)$$

I.21. Sean $R = \{\bar{u}_1, \bar{u}_2, ..., \bar{u}_p\}$ y $S = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_p\}$ dos sistemas de igual número de vectores (de $\mathbb{R}^n$, por ejemplo) cuyos rangos respectivos son r y s. Considérese el sistema $T = \{\bar{w}_i = \bar{u}_i + \bar{v}_i \,/\, i = 1, 2, ..., p\}$. Pruébese que rang $T \leq r + s$.

I.22. Hallar el rango de la siguiente matriz:

$$A = \begin{bmatrix} 1 & 1 & -1 & 1 & -2 & -1 \\ 4 & -5 & 7 & -2 & -4 & -6 \\ 2 & 5 & -8 & 4 & -3 & 1 \\ 3 & -3 & 4 & -1 & -3 & -4 \end{bmatrix}$$

Hallar el rango de la siguiente matriz:

$$A=\begin{bmatrix} 1 & 2 & 0 & -5 & -3 & -2 \\ 3 & 1 & -1 & 3 & 2 & 4 \\ -2 & 0 & 2 & 1 & -1 & 5 \\ 2 & -3 & 1 & 3 & 2 & -1 \\ 1 & -1 & 3 & -1 & -2 & 2 \\ 4 & 3 & -1 & -2 & -1 & 2 \end{bmatrix}$$

I.24. Hallar los valores de α y β para los que el rango de la matriz A es lo más pequeño posible:

$$A=\begin{bmatrix} 1 & 3 & -2 & -1 & 4 \\ -2 & 1 & 1 & 2 & -3 \\ 3 & -4 & 3 & 1 & -2 \\ 3 & 3 & 0 & \alpha & 3 \\ 3 & 2 & -3 & -3 & \beta \end{bmatrix}$$

Hallar el rango de la siguiente matriz compleja:

$$A=\begin{bmatrix} 1 & i & 3i & 3 \\ 2+i & 1 & 1+2i & 4+i \\ -1+i & 1+i & 1+i & -1+i \end{bmatrix}$$

Hallar el rango de la siguiente matriz, de tamaño $n \times n$:

$$A=\begin{bmatrix} x & 0 & 0 & 0 & \cdots & 0 & y \\ y & x & 0 & 0 & \cdots & 0 & 0 \\ 0 & y & x & 0 & \cdots & 0 & 0 \\ 0 & 0 & y & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & x & 0 \\ 0 & 0 & 0 & 0 & \cdots & y & x \end{bmatrix} \text{ para } x \neq 0 \text{ e } y \neq 0$$

Analizar, primero, los casos $n = 3$ y $n = 4$.

I.27. Sea $M = AB$ el producto de las matrices A y B. Compruébese que si A tiene una fila nula, entonces también M tiene una fila nula y que si B tiene una columna nula, entonces M también tiene una columna nula.

I.28. Hallar los números reales x, y, z, u y v para los que se verifica:

$$\begin{bmatrix} x & 2 \\ 0 & y \\ z & 1 \end{bmatrix}\begin{bmatrix} -1 & v \\ u & 0 \end{bmatrix}=\begin{bmatrix} 5 & 1 \\ -3 & 0 \\ 1 & 2 \end{bmatrix}$$

I.29. Sabiendo que $AM = B$, donde A, B y M son matrices cuadradas de igual tamaño, y si A y B son triangulares superiores y no nulas. Se pide:

a) Analizar si M ha de ser triangular superior.

b) Encontrar alguna condición suficiente que permita garantizar que M es triangular superior.

I.30. Sean A y B dos matrices cuadradas de igual tamaño que son simétricas. Hallar una condición necesaria y suficiente para que su producto AB sea simétrica.

I.31. Hallar A^3, siendo A la matriz:

$$A=\begin{bmatrix} 0 & \cos\theta & \operatorname{sen}\theta \\ \cos\theta & 0 & -1 \\ \operatorname{sen}\theta & 1 & 0 \end{bmatrix}$$

I.32. Hallar una matriz A tal que:

$$A^3=\begin{bmatrix} 1 & -7 & 1 \\ 0 & 8 & 38 \\ 0 & 0 & 27 \end{bmatrix}$$

I.33. Hallar A^n, para $n \in \mathbb{N}$, donde:

$$A=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

I.34. Hallar A^p, para $p \geq n$, donde A es la siguiente matriz de tamaño $n \times n$:

$$A=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

I.35. Calcular A^2 y A^3, donde A es la matriz de tamaño $n \times n$:

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

I.36. Hallar todas las matrices de tamaño 4×4 que conmutan con la siguiente matriz A:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

I.37. Una matriz cuadrada M tal que $M^2 = M$ se llama «idempotente». Sabiendo que A y B son matrices cuadradas tales que $A = AB$ y que $B = BA$, compruébese que A y B son idempotentes.

I.38. Se llama *traza* de una matriz $A = [a_{ij}]$, cuadrada de tamaño $n \times n$, a la suma de los elementos de su diagonal; esto es, a:

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

Compruébese que, si A y B son matrices del mismo tamaño $n \times n$, se verifica que:

1. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$
2. $\operatorname{tr}(AA^t) > 0$ si $A \neq O$.

I.39. Considere la familia de las matrices $A(a, h)$, de tamaño $n \times n$, dependientes de los parámetros reales a y h, siguientes:

$$A(a, h) = \begin{bmatrix} ah & a & a & \cdots & a \\ a & ah & a & \cdots & a \\ a & a & ah & \cdots & a \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a & a & a & \cdots & ah \end{bmatrix}$$

a) Compruébese que el producto de $A(a, h)$ por $A(a', h')$ es también una matriz de la familia.

b) Analícese si $A(a, h)$ tiene matriz inversa perteneciente a la familia.

(Indicación: recúrrase a las matrices $M = A(1, 1)$ e I, matriz unidad, expresando $A(a, h)$ como combinación de ellas.)

I.40. Sean A y B dos matrices cuadradas del mismo tamaño. Si A es invertible y A y B conmutan, pruébese que A^{-1} y B también conmutan.

I.41. Hallar la matriz inversa, si existe, de cada una de las siguientes matrices:

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 3 & 7 & 9 \\ 1 & 5 & 1 \end{bmatrix} \quad ; \quad B = \begin{bmatrix} 1 & 1 & 5 \\ 4 & 10 & 16 \\ 2 & 5 & 8 \end{bmatrix}$$

I.42. Hallar la inversa de la matriz:

$$A = \begin{bmatrix} 1 & -1 & 2 & 1 \\ -2 & 3 & -4 & 1 \\ 5 & -8 & 11 & -4 \\ -2 & 3 & -4 & 2 \end{bmatrix}$$

I.43. Hallar las inversas de las matrices:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad ; \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix}$$

I.44. Hallar la inversa de la matriz:

$$A = \begin{bmatrix} -2 & -9 & -8 & 5 \\ -3 & -12 & 10 & 6 \\ 0 & -2 & 2 & 1 \\ -2 & -6 & 5 & 3 \end{bmatrix}$$

I.45. Sea A una matriz cuadrada dada tal que $A^3 = O$. Considérese la familia que forman las matrices:

$$M(\lambda) = I + \lambda A + (\lambda^2/2)A^2 \quad , \text{ para } \lambda \in \mathbb{R}$$

Compruébese que la familia es un grupo abeliano (conmutativo) respecto del producto de matrices; hallar $M(\lambda)^{-1}$.

I.46. Sean A y B dos matrices cuadradas, de tamaño $n \times n$. Se dice que A es *semejante* con B, y se pone $A \sim B$, si existe una matriz regular P tal que $B = P^{-1}AP$. Compruébese que:

1. $(A \sim B) \Rightarrow A^p \sim B^p$ (para $p \in \mathbb{N}$ cualquiera)
2. La semejanza de matrices es una relación de equivalencia para el conjunto $\mathcal{M}_{n\times n}$, de las matrices de tamaño $n \times n$.

I.47. Hallar la matriz inversa de la matriz A:

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a & 1 & 0 & \cdots & 0 \\ a^2 & a & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a^n & a^{n-1} & a^{n-2} & \cdots & 1 \end{bmatrix}$$

I.48. Hallar la inversa de la siguiente matriz compleja:

$$A = \begin{bmatrix} 1 & 2-i & -1+2i \\ 1 & 2 & -2+i \\ -1 & -2+i & 2-2i \end{bmatrix}$$

I.49. Sea A la matriz cuadrada, de tamaño $n \times n$, cuyos elementos son:

$$\text{elemento de lugar } (i, j) \text{ de } A = \begin{cases} a, & \text{si } i = j \\ b, & \text{si } i \neq j \end{cases}$$

Hallar a y b para que se verifique la relación $A^2 = I$.

I.50. Sean A y S dos matrices cuadradas; S es simétrica. Analícese si se verifica que:

a) A^tA es simétrica.
b) A^tSA es simétrica.
c) A antisimétrica $\Rightarrow$ A^2 simétrica.
d) A regular $\Rightarrow$ $A^{-n} = (A^{-1})^n$ $(n \in \mathbb{N})$.

I.51. Se dice que una matriz cuadrada A de elementos complejos es *unitaria* si:

$$A\bar{A}^t = I$$

($\bar{A}$ denota a la matriz conjugada de A, esto es, aquella cuyos elementos son los números complejos conjugados de los respectivos elementos de A). Se pide:

a) Estudiar si la inversa y la traspuesta de una matriz unitaria son matrices unitarias.
b) Si A es unitaria, hallar $|\det A|$.
c) Si A y B son matrices unitarias del mismo tamaño, estudiar si AB es también una matriz unitaria.

I.52. Si A es una matriz cuadrada tal que $A^2 = O$ (matriz nula), hallar $A(I \pm A)^n$, para $n = 2, 3, \ldots$

I.53. Hallar la matriz inversa de la matriz cuadrada de tamaño $n \times n$:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & 3 & \cdots & 3 \\ 1 & 2 & 3 & 4 & \cdots & 4 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 2 & 3 & 4 & \cdots & n \end{bmatrix}$$

I.54. Resolver la ecuación:

$$\begin{vmatrix} x & a & b & c \\ a & x & c & b \\ b & c & x & a \\ c & b & a & x \end{vmatrix} = 0$$

I.55. Resolver la ecuación:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & x & x^2 & x^3 \\ 3 & x+2 & 2x+1 & 3x \\ 3 & 2x+1 & x^2+2x & 3x^2 \end{vmatrix} = 0$$

I.56. Hallar el valor del determinante:

$$\Delta = \begin{vmatrix} \alpha^2 & \alpha\beta & \alpha\beta & \beta^2 \\ \alpha\beta & \alpha^2 & \beta^2 & \alpha\beta \\ \alpha\beta & \beta^2 & \alpha^2 & \alpha\beta \\ \beta^2 & \alpha\beta & \alpha\beta & \alpha^2 \end{vmatrix}$$

I.57. Hallar el determinante de la siguiente matriz, comprobando que no depende de x:

$$A = \begin{bmatrix} \cos(x+a_1) & \cos(x+a_2) & \cos(x+a_3) \\ \operatorname{sen}(x+a_1) & \operatorname{sen}(x+a_2) & \operatorname{sen}(x+a_3) \\ \operatorname{sen}(a_2-a_3) & \operatorname{sen}(a_3-a_1) & \operatorname{sen}(a_1-a_2) \end{bmatrix}$$

Si a_1, a_2 y a_3 pertenecen al intervalo $]-\pi/2, \pi/2[$, hallar la relación entre ellos para que A sea singular.

I.58. Calcular los siguientes determinantes:

$$\Delta_1 = \begin{vmatrix} 3 & -1 & 3 & 4 \\ 4 & -3 & 5 & 6 \\ 4 & 4 & 1 & 3 \\ 3 & 5 & -6 & 7 \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} 3 & 9 & -1 & 2 \\ 1 & 4 & 9 & 0 \\ 2 & 5 & -1 & 1 \\ 6 & 5 & 2 & 1 \end{vmatrix}$$

I.59. Calcular los siguientes determinantes:

$$\Delta_1 = \begin{vmatrix} 9 & 7 & 5 & 2 \\ 4 & -8 & 8 & 3 \\ 6 & -5 & 8 & 4 \\ 7 & 5 & 3 & 7 \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} 10 & 11 & 4 & 1 \\ -1 & 4 & 5 & -1 \\ 5 & 8 & -4 & 4 \\ 2 & 4 & 1 & 1 \end{vmatrix}$$

I.60. Calcular el siguiente determinante:

$$\Delta = \begin{vmatrix} 1 & 3 & 5 & 2 & 2 \\ 3 & 6 & 9 & 5 & 1 \\ 6 & 5 & 7 & 4 & 2 \\ 1 & 6 & 8 & 5 & 0 \\ 1 & 4 & 6 & 3 & 1 \end{vmatrix}$$

I.61. Calcular el determinante:

$$\Delta = \begin{vmatrix} n & 1 & 1 & 1 & \cdots & 1 \\ n & 2 & 1 & 1 & \cdots & 1 \\ n & 1 & 3 & 1 & \cdots & 1 \\ n & 1 & 1 & 4 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ n & 1 & 1 & 1 & \cdots & n \end{vmatrix}$$

I.62. Sean A y B dos matrices cuadradas de igual tamaño. Hállese el determinante de la siguiente matriz M en función de los determinantes de $A+B$ y de $A-B$:

$$M = \left[\begin{array}{c|c} A & B \\ \hline B & A \end{array}\right]$$

I.63. Calcular el determinante de orden n:

$$\Delta_n = \begin{vmatrix} x_1+y_1 & x_1+y_2 & \cdots & x_1+y_n \\ x_2+y_1 & x_2+y_2 & \cdots & x_2+y_n \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ x_n+y_1 & x_n+y_2 & \cdots & x_n+y_n \end{vmatrix}$$

I.64. Sea $A = [a_{ij}]$ una matriz cuadrada de tamaño $n \times n$ y considérese la nueva matriz $B = [a_{ij} + k]$, que se obtiene de sumar un mismo escalar k a cada uno de los elementos de A. Hállese $\det B$ en función de k, de los adjuntos α_{ij} de los elementos de A y de $\det A$.

I.65. Se considera el siguiente determinante de orden n:

$$\Delta_n = \begin{vmatrix} 1+a_1 & a_2 & a_3 & \vdots & \vdots & a_n \\ a_1 & 1+a_2 & a_3 & \vdots & \vdots & a_n \\ a_1 & a_2 & 1+a_3 & \vdots & \vdots & a_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_1 & a_2 & a_3 & \vdots & \vdots & 1+a_n \end{vmatrix}$$

Obtener la relación que liga a Δ_n con Δ_{n-1}. Calcular Δ_n.

I.66. Sea A una matriz cuadrada, de tamaño $n \times n$, antisimétrica ($A^t = -A$). Demuéstrese que si n es impar, entonces $\det A = 0$.

I.67. Calcular los determinantes de orden n:

$$\Delta_1 = \begin{vmatrix} 1 & 2 & 3 & \cdots & n \\ -1 & 0 & 3 & \cdots & n \\ -1 & -2 & 0 & \cdots & n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & -2 & -3 & \cdots & 0 \end{vmatrix}$$

$$\Delta_2 = \begin{vmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ 1 & x_1 + a_1 & x_2 & \cdots & x_n \\ 1 & x_1 & x_2 + a_2 & \cdots & x_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_1 & x_2 & \cdots & x_n + a_n \end{vmatrix}$$

I.68. Considérese el determinante de orden n:

$$\Delta_n = \begin{vmatrix} a + x_1 & x_2 & x_3 & \cdots & x_n \\ x_1 & a + x_2 & x_3 & \cdots & x_n \\ x_1 & x_2 & a + x_3 & \cdots & x_n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_1 & x_2 & x_3 & \cdots & a + x_n \end{vmatrix}$$

Obtener Δ_n en función de Δ_{n-1} y, de ello, hallar Δ_n.

I.69. Obtener la relación que liga a Δ_n con Δ_{n-1}, siendo Δ_n el siguiente determinante de orden n:

$$\Delta_n = \begin{vmatrix} 0 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 0 & x & x & \cdots & x \\ 1 & x & 0 & x & \cdots & x \\ 1 & x & x & 0 & \cdots & x \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x & x & x & \cdots & 0 \end{vmatrix}$$

Calcular Δ_n.

I.70. Discutir el siguiente sistema de ecuaciones, resolviéndolo siempre que sea posible:

$$\left.\begin{aligned} ax + by + z &= 1 \\ x + aby + z &= b \\ x + by + az &= 1 \end{aligned}\right\} \quad (a, b \in \mathbb{R} \text{ son parámetros})$$

I.71. Resolver el siguiente sistema de ecuaciones:

$$\left.\begin{aligned} \alpha x_1 + \alpha x_2 + \cdots + \alpha x_{n-1} + \beta x_n &= a_n \\ \alpha x_1 + \alpha x_2 + \cdots + \beta x_{n-1} + \alpha x_n &= a_{n-1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha x_1 + \beta x_2 + \cdots + \alpha x_{n-1} + \alpha x_n &= a_2 \\ \beta x_1 + \alpha x_2 + \cdots + \alpha x_{n-1} + \alpha x_n &= a_1 \end{aligned}\right\} \quad (\alpha \neq 0 \neq \beta)$$

I.72. Discutir y resolver, en caso de compatibilidad, el sistema:

$$\left.\begin{aligned} ax + by + z &= 1 \\ ax + y + bz &= 1 \\ ax + y + z &= b \end{aligned}\right\} \quad (a, b \in \mathbb{R} \text{ son parámetros})$$

I.73. Discutir, según los valores de los parámetros a y b, y resolver, en su caso, el siguiente sistema de ecuaciones:

$$\left.\begin{aligned} x_1 - x_3 + x_4 &= 0 \\ x_1 - x_2 + x_3 &= -1 \\ ax_1 + x_3 + x_4 &= 0 \\ bx_1 - x_2 - x_4 &= -1 \\ x_2 + bx_3 + x_4 &= a \end{aligned}\right\}$$

I.74. Discutir, según los valores de los parámetros a y b, el siguiente sistema de ecuaciones. Resolver el sistema para $a = 3$ y $b = 2$:

$$\left.\begin{aligned} 2x_1 - x_2 - 3x_4 - x_5 &= 2 \\ -x_1 + x_2 + x_3 + 2x_4 + 2x_5 &= 3 \\ x_1 + x_2 + 3x_3 + 2x_5 &= 6 \\ x_2 + 2x_3 + ax_4 + x_5 &= b \end{aligned}\right\}$$

I.75. Discutir, según los valores de los parámetros reales a y b, el siguiente sistema de ecuaciones ($n+1$ ecuaciones con $n+1$ incógnitas). Resolver el sistema cuando sea compatible:

$$\left.\begin{aligned} x_1 + ax_{n+1} &= a \\ x_2 + ax_{n+1} &= a \\ &\cdots\cdots\cdots \\ &\cdots\cdots\cdots \\ x_n + ax_{n+1} &= a \\ a(x_1 + x_2 + \cdots + x_n) + x_{n+1} &= b \end{aligned}\right\}$$

I.76. Resolver el sistema de ecuaciones lineales:

$$\left.\begin{aligned} ix_1 + 2x_2 + (1+i)x_3 + (1-i)x_4 &= -2+i \\ (1-i)x_1 + ix_2 + 2x_3 - 2ix_4 &= 2 \\ -2x_1 + (1+i)x_2 + ix_3 - (2-i)x_4 &= 2-3i \\ 3ix_1 - 2ix_2 + (1-2i)x_3 - x_4 &= 0 \end{aligned}\right\}$$

I.77. Sean A y B matrices de tamaños $m \times n$ y $m \times 1$; supóngase que $r = \text{rang } A = \text{rang } (A|B) < m, n$. Considérese el sistema de ecuaciones $AX = B$, donde X es la matriz de incógnitas, de tamaño $n \times 1$. Se descomponen A y B y X en bloques del siguiente modo:

$$A = \left[\begin{array}{c|c} A_1 & A_2 \\ \hline A_3 & \end{array}\right] \quad ; \quad B = \left[\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right] \quad ; \quad X = \left[\begin{array}{c} X_1 \\ \hline X_2 \end{array}\right]$$

donde A_1 es una matriz regular de tamaño $r \times r$; B_1 y X_1 son de tamaño $r \times 1$. Se pide:

a) Obtener, en función de dichos bloques las soluciones del sistema.

b) Obtener las soluciones del sistema:

$$\left.\begin{aligned} x + 2y + z + 9u + 15v &= -2 \\ 2x + 3y + z + 17u + 21v &= 1 \\ -x + y + z - 8u + 7v &= 1 \\ x + 13y + 8z + 12u + 94v &= -1 \end{aligned}\right\}$$

SOLUCIONES

I.1. *a*) $x = 1, y = -1, z = 2$
b) $x = 2, y = -3, z = 1$

I.2. *a*) El sistema es incompatible.
b) $x = 2 - z, y = 1 + 2z, z$ cualquiera.

I.3. *a*) $\alpha = -10; x = 3 + 2z, y = -1 - z, z$ cualquiera.
b) $\alpha = -1; x = 2, y = -1, z = 0$

I.4. *a*) $x_1 = 1, x_2 = 1, x_3 = 1, x_4 = 2$
b) $x_1 = 2, x_2 = 0, x_3 = -1, x_4 = 1$

I.5. $\alpha = 4; x = 8 - 18z + 11u, y = -3 + 7z - 4u$

I.6. $x = 3 - u - 2v, y = -1 + 2u + v, z = 2 - v$

I.7. $x = 4, y = -2, z = 3, u = 1$

I.8. Si $\alpha = 0$, el sistema es incompatible.
Si $\alpha = 5$, hay infinitas soluciones:

$$x = 5z \quad , \quad y = 1 - 4z$$

Si $\alpha \neq 0$ y $\alpha \neq 5$, hay solución única.

$$x = 1 \quad , \quad y = \frac{1}{\alpha} \quad , \quad z = \frac{1}{\alpha}$$

I.9. Reducir el sistema S a uno escalonado reducido y analizar éste.

I.10. $x_1 = -1 + x_5, x_2 = 0, x_3 = x_5, x_4 = 1 - x_5$

I.11. Para $\alpha = -2$ el sistema es incompatible. Para $\alpha = 1$, el sistema admite infinitas soluciones: $x = 1 - y - z$, y cualquiera, z cualquiera. Para $\alpha \neq -2$ y $\alpha \neq 1$, el sistema tiene solución única:

$$x = -\frac{\alpha + 1}{\alpha + 2} \quad , \quad y = \frac{1}{\alpha + 2} \quad , \quad z = \frac{(\alpha + 1)^2}{\alpha + 2}$$

I.12. Si, por ejemplo:

$$\bar{u}_1 = (1, 0, 0) \quad , \quad \bar{u}_2 = (0, 1, 0)$$
$$\bar{v}_1 = (1, 1, 1) \quad , \quad \bar{v}_2 = (1, 1, -1)$$
$$p = h = 2 \quad ; \quad \bar{v}_1 + \bar{v}_2 = 2\bar{u}_1 + 2\bar{u}_2$$

I.13. *a*) Linealmente dependiente.
b) Linealmente independiente.

I.14. Son todas distintas.

I.15. Puede no serlo, como en el ejemplo:

$$p = q = 2 \quad , \quad \bar{v}_1 = (1, 0, 0) \quad , \quad \bar{v}_2 = (0, 1, 0)$$
$$\bar{u}_1 = (1, 0, 1) \quad , \quad \bar{u}_2 = (0, 1, 1)$$

aquí es $\bar{u}_1 + \bar{v}_2 = \bar{v}_1 + \bar{u}_2$

I.16. Supóngase que hay una combinación lineal de las $\bar{u}_i$ que es nula; recurriendo a los componentes se obtiene un sistema de ecuaciones lineales homogéneo con más incógnitas (los escalares) que ecuaciones, que tendrá, pues, solución no nula.

I.17. *a*) Dependiente: $i\bar{u} - i\bar{v} = \bar{w}$:
b) Independiente.

I.18. No existen: los $\bar{u}$, $\bar{v}$ y $\bar{w}$ no son independientes y los $\bar{x}, \bar{y}, \bar{z}$ dependen de ellos, luego son dependientes siempre.

I.19. Si $h \neq 1 - n$ y $h \neq 1$, el rango es n; si $h = 1 - n$, el rango es $n - 1$; si $h = 1$, el rango vale 1.

I.20. rang $S = 4$.

I.21. Sean $\bar{u}_{i_1}, ..., \bar{u}_{i_r}$ y $\bar{v}_{j_1}, ..., \bar{v}_{j_s}$ sistemas independientes de R y S. Todos los vectores de T dependen linealmente de los vectores de $\{\bar{u}_{i_1}, ..., \bar{u}_{i_r}\} \cup \{\bar{v}_{j_1}, ..., \bar{v}_{j_s}\}$ y aquí hay no más de $r + s$ vectores.

I.22. El rango es 3.

I.23. El rango es 4.

I.24. $\alpha = 1, \beta = 7$; el rango mínimo es 3.

I.25. rang $A = 2$.

I.26. Para $n = 3$, rang $A = 3$ si $a \neq b$ y rang $A = 2$ si $a = b$. Para $n = 4$, rang $A = 4$ si $a \neq \pm b$ y rang $A = 3$ si $a = \pm b$. Para n par rang $A = n$ si $a \neq \pm b$ y rang $A = n - 1$ si $a = \pm b$. Para n impar, rang $A = n$ si $a \neq b$ y rang $A = n - 1$ si $a = b$.

I.27. Si $a_{ih}=0$, $\forall h$, entonces.

$$m_{ij}=\sum_h a_{ih}b_{hj}=\sum_h 0b_{hj}=0,\ \forall j$$

Para columnas se procede igualmente.

I.28. $x=1,\ y=-1,\ z=2,\ u=3,\ v=1$

I.29. *a*) Puede no serlo; así ocurre con:

$$A=\begin{bmatrix}1 & 2\\ 0 & 0\end{bmatrix}\quad \text{y}\quad M=\begin{bmatrix}3 & -2\\ 1 & 3\end{bmatrix}$$

b) Que A sea invertible.

I.30. Que A y B conmuten.

I.31. $A^3=O$.

I.32. $A=\begin{bmatrix}1 & -1 & 1\\ 0 & 2 & 2\\ 0 & 0 & 3\end{bmatrix}$

I.33. $A^n=\begin{bmatrix}1 & n & s\\ 0 & 1 & n\\ 0 & 0 & 1\end{bmatrix}$, con $s=\dfrac{n(n-1)}{2}$

I.34. $A^p=O$.

I.35. $A^2=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ 2 & 1 & 0 & \cdots & 0\\ 3 & 2 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ n & n-1 & n-2 & \cdots & 1\end{bmatrix}$

$$A^3=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ 3 & 1 & 0 & \cdots & 0\\ 6 & 3 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ s_n & s_{n-1} & s_{n-2} & \cdots & 1\end{bmatrix}$$

con $s_i=\dfrac{i(i-1)}{2}$

I.36. $\begin{bmatrix}a & 0 & 0 & 0\\ b & a & 0 & 0\\ c & b & a & 0\\ d & c & b & a\end{bmatrix}$

para a, b, c, d escalares cualesquiera.

I.37. $A^2=(AB)(AB)=A(BA)B=ABB=(AB)B=AB=A$
igual se procede con B^2.

I.38. 1. Ambos trazos valen $\sum_{i,j} a_{ij}b_{ji}$.

2. $\operatorname{tr}(AA^t)=\sum_{i,j}(a_{ij})^2>0$.

I.39. *a*) El producto vale $\alpha M+\beta I$, donde

$$\alpha=aa'(n+h+h'-2)\ ,\ \beta=aa'(h-1)(h'-1)$$

b) $A^{-1}=A(a_1,\ h_1)$, con

$$h_1=2-h-n\quad \text{y}\quad a_1=\frac{1}{a(h-1)(1-h-n)}$$

(para $h\neq 1$, $h\neq 1-n$ y $a\neq 0$).

I.40. $AB=BA\ \Rightarrow\ A^{-1}(AB)A^{-1}=A^{-1}(BA)A^{-1}\ \Rightarrow$
$\Rightarrow\ (A^{-1}A)(BA^{-1})=(A^{-1}B)(AA^{-1})\ \Rightarrow$
$\Rightarrow\ IBA^{-1}=A^{-1}BI\ \Rightarrow\ BA^{-1}=A^{-1}B$

I.41. $A^{-1}=\begin{bmatrix}-19 & 3 & 11\\ 3 & -1/2 & -3/2\\ 4 & -1/2 & -5/2\end{bmatrix}$; no existe B^{-1}.

I.42. $A^{-1}=\begin{bmatrix}1 & -1 & -2 & -4\\ 2 & 4 & 0 & -3\\ 1 & 3 & 1 & 0\\ 0 & -1 & 0 & 1\end{bmatrix}$

I.43. $A^{-1}=\begin{bmatrix}1 & 0 & 0 & 0\\ -1 & 1 & 0 & 0\\ 0 & -1 & 1 & 0\\ 0 & 0 & -1 & 1\end{bmatrix}$

$$B^{-1}=\begin{bmatrix}1&0&0&0\\-2&1&0&0\\1&-2&1&0\\0&1&-2&1\end{bmatrix}$$

I.44. $A^{-1}=\begin{bmatrix}0&1&0&-2\\1&-2&1&2\\0&-2&3&-3\\2&0&-3&-2\end{bmatrix}$

I.45. $M(0)=I$, $M(\lambda)\cdot M(\mu)=M(\mu)\cdot M(\lambda)=M(\lambda+\mu)$; $M(\lambda)^{-1}=M(-\lambda)$.

I.46. 1. $B^p=BB\cdots B=$
$=(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)=$
$=P^{-1}AA\cdots AP=P^{-1}A^pP$.

2. $A=I^{-1}AI\Rightarrow A\sim A$;
$A\sim B\Rightarrow B=P^{-1}AP\Rightarrow A=PBP^{-1}\Rightarrow B\sim A$;
$(A\sim B$ y $B\sim C)\Rightarrow$
$\Rightarrow (B=P^{-1}AP$ y $C=Q^{-1}BQ)\Rightarrow$
$\Rightarrow C=(PQ)^{-1}A(PQ)\Rightarrow A\sim C$

I.47. $A^{-1}=\begin{bmatrix}1&0&0&\cdots&0\\-a&1&0&\cdots&0\\0&-a&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&1\end{bmatrix}$

I.48. $A^{-1}=\begin{bmatrix}-i&1+2i&i\\1&-i&1-i\\1&0&1\end{bmatrix}$

I.49. Póngase A en la forma $A=aI=bB$. Solución:

$$a=\pm(n-2):n \quad \text{y} \quad b=\mp 2:n$$

I.50. Los cuatro apartados son ciertos.

I.51. *a*) Ambos son unitarios.
b) $|\det A|=1$.
c) AB es unitaria.

I.52. $A(I\pm A)^n=A$.

I.53. $A^{-1}=\left[\begin{array}{cccc:cc}2&-1&0&0&0&0\\-1&2&-1&0&0&0\\0&-1&2&-1&0&0\\0&0&-1&2&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&0&0&2&-1\\0&0&0&0&-1&1\end{array}\right]$.

I.54. Tiene 4 soluciones, que son:

$$x_1=-a-b-c \;;\; x_2=a+b-c$$
$$x_3=-a+b+c \;;\; x_4=a-b+c$$

I.55. Sólo tiene la raíz $x=1$ (de sexto orden de multiplicidad).

I.56. $(\alpha^2-\beta^2)^4$.

I.57. Desarrollando por la última fila:

$$\det A=-\operatorname{sen}^2(a_2-a_3)-\operatorname{sen}^2(a_3-a_1)-$$
$$-\operatorname{sen}^2(a_1-a_2)$$

La condición es $a_1=a_2=a_3$.

I.58. $\Delta_1=130$; $\Delta_2=61$.

I.59. $\Delta_1=100$; $\Delta_2=150$.

I.60. $\Delta=-5$.

I.61. $\Delta=n!$

I.62. $\det M=[\det(A+B)]\cdot[\det(A-B)]$.

I.63. $\Delta_1=x_1+y_1$; $\Delta_2=-(x_1-x_2)(y_1-y_2)$
$\Delta_n=0$ para $n\geqslant 3$.

I.64. $\det B=\det A+k\sum_{i,j=1}^{n}\alpha_{ij}$

I.65. $\Delta_n=\Delta_{n-1}+a_n$; $\Delta_n=1+a_1+a_2+\cdots+a_n$

I.66. $\det A=\det(-A^t)=(-1)^n\det A^t=(-1)^n\det A=$
$=-\det A$, luego $\det A=0$.

I.67. $\Delta_1 = n! \quad ; \quad \Delta_2 = a_1 a_2 \cdots a_n$

I.68. $\Delta_n = a\Delta_{n-1} + a^n x_n$
$\Delta_n = a^{n-1}(a + x_1 + x_2 + \cdots + x_n)$

I.69. $\Delta_n = (-1)^{n+1} x^{n-2} - x\Delta_{n-1}$
$\Delta_n = (-1)^{n+1}(n-1)x^{n-2}$

I.70. Para $b = 0$ y $a = 1$, sistema incompatible. Para $b = 0$ y $a \neq 1$, sistema incompatible. Para $a = 1$ y $b = 1$, infinitas soluciones: $x = \alpha$, $y = \beta$, $z = 1 - \alpha - \beta$. Para $a = 1$ y $b \neq 1$, sistema incompatible. Para $a = -2$ y $b = -2$, infinitas soluciones: $x = 3\alpha$, $y = 1/2(\alpha - 1)$, $z = \alpha$. Para $a = -2$ y $b \neq -2$, sistema incompatible.

I.71. $$x_i = \frac{\alpha\left(\sum_{j=1}^{n} a_j\right) - [\alpha(n-1) + \beta]a_i}{(\alpha - \beta)[\alpha(n-1) + \beta]}$$

I.72. Para $b = 1$ y $a = 0$, compatible indeterminado; las soluciones son $(x, y, z) = (\alpha, \beta, 1 - \beta)$. Para $b = 1$ y $a \neq 0$, compatible indeterminado; $(x, y, z) = (\alpha, \beta, 1 - \beta - a\alpha)$. Para $b \neq 1$ y $a = 0$, sistema compatible determinado; $(x, y, z) = ((b+2):a, -1, -1)$. Para $a = 0$ y $b = -2$, sistema compatible indeterminado; $(x, y, z) = (\alpha, -1, -1)$. Para $b \neq -1$, $b \neq -2$ y $a = 0$, sistema incompatible.

I.73. Si $b \neq -2$, el sistema tiene solución única:

$$x_1 = \frac{-2}{b+2} \quad , \quad x_2 = \frac{a+b-1}{b+2} \quad , \quad x_3 = \frac{a-1}{b+2}$$

$$x_4 = \frac{a+1}{b+2}$$

Si $b = -2$, el sistema es incompatible

I.74. Para $a \neq 1$, hay infinitas soluciones. Para $a = 1$ y $b \neq 1$ no hay solución. Para $a = 1$ y $b = 1$, hay infinitas soluciones. En el caso $a = 3$, $b = 2$, las soluciones son, para $\lambda \in \mathbb{R}$ cualquiera:

$$x_1 = \lambda + 8 \quad , \quad x_2 = -2\lambda - 3 \quad , \quad x_3 = \lambda$$

$$x_4 = \frac{1}{2} \quad , \quad x_5 = \frac{7}{2}$$

I.75. Si $a = 0$, hay solución única $x_1 = x_2 = \cdots = x_n = 0$, $x_{n+1} = b$. Si $a \neq 0$ y $a \neq \pm\sqrt{1/n}$, hay solución única, que es la:

$$x_1 = x_2 = \cdots = x_n = \frac{a(1-b)}{1 - a^2 n}$$

y

$$x_n = \frac{b - a^2 n}{1 - a^2 n}$$

Si es $a = \pm\sqrt{1 n}$ y $b \neq 1$, no hay solución. Si $a = \pm\sqrt{1/n}$ y $b = 1$, hay infinitas soluciones, que son (para $\lambda \in \mathbb{R}$ cualquiera):

$$x_1 = x_2 = \cdots = x_n = \frac{\lambda}{\pm\sqrt{n}} \quad \text{y} \quad x_{n+1} = 1 - \lambda$$

I.76. $x_1 = i \quad , \quad x_2 = 1 - i \quad , \quad x_3 = 2i \quad , \quad x_4 = -1.$

I.77. *a*) X_2 cualquiera, $X_1 = A_1^{-1}(B_1 - A_2 X_2)$.
b) u y v cualesquiera:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & -1 \\ 1 & 1 & \end{bmatrix}^{-1} \left\{ \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} - \right.$$

$$\left. - \begin{bmatrix} 9u + 15v \\ 17u + 21v \\ -8u + 7v \end{bmatrix} \right\} = \begin{bmatrix} -6 - 9u - 2v \\ 9 + u - 3v \\ -14 - 2u - 5v \end{bmatrix}$$

PARTE II

AÚN cuando ésta ha resultado ser la segunda parte, a causa de haber otra que la precede, no deja por ello de ser la primera, ya que es la más importante y principal de cuantas componen esta obra. Empero no se ha de alarmar nadie por ello, que aquello que tiene mayor interés no suele ser lo más difícil y, aquí, si bien se mira, en el fondo de cuanto se dice sólo hay un concepto básico, sobre el que reposa todo el negocio, el concepto de «lo lineal». A no dudarlo, esta idea nos ha de ser familiar y conocida de antiguo, pues nos fue presentada por nuestros maestros, cuando, siendo aún tiernos escolares, nos hablaron de la «regla de tres».

5. Espacios vectoriales.
6. Aplicaciones lineales. Ejercicios y problemas.

CAPITULO

5

Espacios vectoriales

A estas alturas nos es ya imprescindible abordar el tema de los vectores con mayor amplitud; se hace necesario generalizar la noción de «vector de n componentes» que hemos venido manejando hasta ahora en la primera parte o «Iniciación: linealidad y rango». El modelo a generalizar es el conjunto de todos los vectores de K^n (n número natural fijo; K cuerpo, por ejemplo, $K = \mathbb{R}$), junto con las operaciones «suma» y «producto por escalar». Las propiedades de las que goza este modelo concreto, que coinciden con las de otros muchos ejemplos, se tomarán como punto de partida, es decir, formarán la axiomática de los espacios vectoriales. Siempre que se cumplan estos axiomas, se dirá que la estructura en cuestión es un espacio vectorial; aquí nada importa cual sea la naturaleza de los vectores que se consideren ni el modo de definir las operaciones, con tal de que se verifiquen los axiomas, que es el único requisito que en esto se exige.

α ESPACIOS, SUBESPACIOS Y COMBINACIONES LINEALES

5.1. CONCEPTO DE ESPACIO VECTORIAL

[052]

Sea V un conjunto dado, a cuyos elementos llamaremos *vectores*. Se considera también un cuerpo K (en particular $K = \mathbb{R}$ y $K = \mathbb{C}$), a cuyos elementos llamaremos *escalares*. Se dice que V es un *espacio vectorial* sobre K si se dispone de las siguientes operaciones «suma» y «producto por escalar» y se satisfacen las propiedades que abajo se detallan:

1.° Una operación $(+)$ interna en V (suma de vectores) tal que se cumplen los siguientes axiomas (de la suma):
- $(\bar{u} + \bar{v}) + \bar{w} = \bar{u} + (\bar{v} + \bar{w})$ $\forall\, \bar{u}, \bar{v}, \bar{w} \in V$ (asociatividad) (S.1)
- $\bar{u} + \bar{v} = \bar{v} + \bar{u}$ $\forall\, \bar{u}, \bar{v} \in V$ (conmutabilidad) (S.2)
- $\exists \bar{o} \in V$ (*vector nulo*) tal que $\bar{u} + \bar{o} = \bar{u}$ para cualquier $\bar{u} \in V$ (S.3)
- Para cada $\bar{u} \in V$ existe $-\bar{u} \in V$ tal que $\bar{u} + (-\bar{u}) = \bar{o}$ (S.4)
  (cada vector $\bar{u} \in V$ tiene un *vector opuesto* $-\bar{u} \in V$)

2.° Una operación externa (producto por escalar), que a cada pareja $\lambda \in K$, $\bar{u} \in V$ asocia un vector $\lambda\bar{u}$, tal que se cumplen los siguientes axiomas (del producto por escalar):
- $\lambda(\bar{u} + \bar{v}) = \lambda\bar{u} + \lambda\bar{v}$ $\forall\, \bar{u}, \bar{v} \in V$ $\forall\, \lambda \in K$ (P.1)
- $(\lambda + \mu)\bar{u} = \lambda\bar{u} + \mu\bar{u}$ $\forall\, \bar{u} \in V$ $\forall\, \lambda, \mu \in K$ (P.2)
- $\lambda(\mu\bar{u}) = (\lambda\mu)\bar{u}$ $\forall\, \bar{u} \in V$ $\forall\, \lambda, u \in K$ (P.3)
- $1\bar{u} = \bar{u}$ (1 = unidad de K) $\forall\, \bar{u} \in V$ (P.4)

OBSERVACIONES

- Usualmente, el cuerpo K será para nosotros $K = \mathbb{R}$ o $K = \mathbb{C}$. Si es $K = \mathbb{R}$, el espacio vectorial se llama real; si $K = \mathbb{C}$. el espacio vectorial se llama complejo.
- Los cuatro axiomas que afectan sólo a la suma (los englobados en el apartado 1), se pueden resumir diciendo que $(V, +)$ es un grupo conmutativo. Los otros cuatro axiomas (los englobados en el apartado 2) regulan el modo de «actuar» los escalares del cuerpo K sobre el grupo $(V, +)$.
- Por ser $(V, +)$ un grupo, se verifica que el vector nulo $\bar{o}$ es único, que cada vector $\bar{u}$ tiene un solo opuesto $-\bar{u}$, que si es $\bar{u} + \bar{v} = \bar{u} + \bar{w}$ entonces es $\bar{v} = \bar{w}$. Como en todo grupo, también aquí se pone $\bar{u} - \bar{v}$ en lugar de $\bar{u} + (-\bar{v})$.

EJEMPLOS

Además del espacio vectorial K^n (véase [011] y [012]), que nos ha servido de modelo, conviene citar a los siguientes:

1. El conjunto $V = \mathcal{F}(C, K)$, de las funciones definidas en un conjunto C y con valores en el cuerpo K, es un espacio vectorial sobre K respecto de las siguientes operaciones: dadas $f, g \in V$ y $\lambda \in K$, se llaman $f + g$ y λf a las funciones de C en K definidas por:

$$\begin{cases} f+g & \text{es tal que} \quad (f+g)(x) = f(x) + g(x) \quad \forall x \in C \\ \lambda f & \text{es tal que} \quad (\lambda f)(x) = \lambda f(x) \quad \forall x \in C \end{cases}$$

2. El conjunto $V = K[x]$, de los polinomios con coeficientes en K y con una indeterminada x, es un espacio vectorial sobre K respecto de las operaciones usuales de suma de polinomios y producto de un escalar por un polinomio.
3. El conjunto $V = \mathcal{M}_{m\times n}$, de las matrices de tamaño $m \times n$ (cuyos elementos pertenecen a un cuerpo K) es un espacio vectorial respecto de la suma de matrices y del producto de un escalar por una matriz (véase [022])

CONSECUENCIAS DE LOS AXIOMAS

[053]

En un espacio vectorial V sobre un cuerpo K, para cualesquiera que sean $\bar{u}, \bar{v} \in V$ y $\lambda, \mu \in V$, se verifica que:

1. $\lambda\bar{o} = \bar{o}$
2. $0\bar{u} = \bar{o}$
3. $\lambda\bar{u} = \bar{o} \Rightarrow [\lambda = 0 \text{ ó } \bar{u} = \bar{o}]$
4. $[\lambda\bar{u} = \mu\bar{u} \text{ y } \bar{u} \neq \bar{o}] \Rightarrow \lambda = \mu$
5. $[\lambda\bar{u} = \lambda\bar{v} \text{ y } \lambda \neq 0] \Rightarrow \bar{u} = \bar{v}$
6. $(-\lambda)\bar{u} = \lambda(-\bar{u}) = -\lambda\bar{u}$

DEMOSTRACION

1. $\lambda\bar{o} = \lambda(\bar{o} + \bar{o}) = \lambda\bar{o} + \lambda\bar{o}$; restando $\lambda\bar{o}$, se obtiene $\bar{o} = \lambda\bar{o}$.
2. $0\bar{u} = (0 + 0)\bar{u} = 0\bar{u} + 0\bar{u}$; restando $0\bar{u}$, se obtiene $\bar{o} = 0\bar{u}$.
3. Si fuese $\lambda = 0$, la propiedad sería cierta. Si es $\lambda \neq 0$, entonces existe λ^{-1}; multiplicando por λ^{-2}, se obtiene $\lambda^{-2}(\lambda\bar{u}) = \lambda^{-1}\bar{o}$, o sea, $(\lambda^{-1}\lambda)\bar{u} = \bar{o}$, esto es $1\bar{u} = \bar{o}$, luego $\bar{u} = \bar{o}$.
4. Se sabe que $\lambda\bar{u} - \mu\bar{u} = \bar{o}$, o sea, $(\lambda - \mu)\bar{u} = \bar{o}$; como es $\bar{u} \neq \bar{o}$, según la propiedad anterior ha de ser $\lambda - \mu = 0$; es decir, $\lambda = \mu$.
5. Se sabe que $\lambda\bar{u} - \lambda\bar{v} = \bar{o}$, o sea, $\lambda(\bar{u} - \bar{v}) = \bar{o}$; como es $\lambda \neq 0$, según la propiedad anterior ha de ser $\bar{u} - \bar{v} = \bar{o}$, es decir, $\bar{u} = \bar{v}$.
6. $(-\lambda)\bar{u} + \lambda\bar{u} = (-\lambda + \lambda)\bar{u} = 0\bar{u} = \bar{o}$, luego $(-\lambda)\bar{u} = -\lambda\bar{u}$. Por otra parte, $\lambda(-\bar{u}) + \lambda\bar{u} = \lambda(-\bar{u} + \bar{u}) = \lambda\bar{o} = \bar{o}$, luego $\lambda(-\bar{u}) = -\lambda\bar{u}$.

5.2. SUBESPACIOS VECTORIALES

Se llaman subespacios de un espacio vectorial V a aquellos subconjuntos de V que son, a su vez, espacios vectoriales respecto de las mismas operaciones de V. Al concretar y desarrollar lo aquí dicho, obtenemos:

DEFINICION Y CARACTERIZACION DE LOS SUBESPACIOS

[054]

Sea V un espacio vectorial sobre K (cuerpo) y sea U un subconjunto de V.

1. Definición.—Se dice que U es un subespacio vectorial de V si las operaciones de V son, también, operaciones para U y, con ellas, U es un espacio vectorial sobre K.
2. Caracterización.—Se verifica que U es un subespacio vectorial de V si, siendo $U \neq \emptyset$, se cumplen las dos condiciones *a*) y *b*):

   a) $\bar{u}, \bar{v} \in U \Rightarrow \bar{u} + \bar{v} \in U$
   b) $\lambda \in K, \bar{u} \in U \Rightarrow \lambda\bar{u} \in U$

Estas condiciones *a*) y *b*) se pueden sustituir, ambas, por la condición *c*):

c) $[\bar{u}, \bar{v} \in U \text{ y } \lambda, \mu \in K] \Rightarrow \lambda\bar{u} + \mu\bar{v} \in U$

Se dice que el vector $\lambda\bar{u} + \mu\bar{v}$ (para cualesquiera que sean los escalares λ y μ) es una *combinación lineal* de los vectores $\bar{u}$ y $\bar{v}$.

DEMOSTRACION

- Es evidente que $1 \Rightarrow 2$ ya que $U \neq \emptyset$, pues $\bar{o} \in K$, y las condiciones *a*) y *b*) significan que las operaciones de V son operaciones para U.
- También se cumple que $2 \Rightarrow 1$, ya que para U se verifican entonces todos los axiomas de espacio vectorial; en efecto: Las condiciones *a*) y *b*) significan que las operaciones de V son operaciones para U. Los axiomas S.1, S.2, P.1, P.2, P.3 y P.4 de espacio vectorial (ver [052]) se verifican en U, ya que se verifican en V. El axioma S.3 se verifica ya que, como $U \neq \emptyset$, existe $\bar{u}_0 \in U$ y, por tanto, $0\bar{u}_0 \in U$, es decir, $\bar{o} \in U$. Finalmente, también se verifica el axioma S.4 ya que, si $\bar{u} \in U$, entonces $(-1)\bar{u} \in U$, es decir, $-\bar{u} \in U$.
- Para concluir, comprobemos que las dos condiciones *a*) y *b*) equivalen a la condición *c*). Si se verifican *a*) y *b*), entonces $\lambda\bar{u} \in U$ y $\mu\bar{v} \in U$ (por *a*) y de aquí se infiere que $\lambda\bar{u} + \mu\bar{v} \in U$ (por *b*). Si se verifica *c*), tomando $\lambda = \mu = 1$ se obtiene *a*) y tomando $\mu = 0$ se obtiene *b*).

OBSERVACIONES

1. El vector nulo $\bar{o}$ pertenece a todos los subespacios de un espacio V.
2. Un espacio vectorial V tiene como subespacios, entre otros posibles, al conjunto $O = \{\bar{o}\}$, formado sólo por el vector nulo, que se llamará subespacio nulo. El propio espacio V es un subespacio de sí mismo. Los demás subespacios de V, distintos de O y V, se llaman subespacios propios.

EJEMPLOS

1. El conjunto $U = \{(x, y, z) \in \mathbb{R}^3 / 3x - 2y + 4z = 0\}$ es un subespacio vectorial del espacio $\mathbb{R}^3$.
2. El conjunto de los polinomios complejos cuyo grado es menor o igual que 5 es un subespacio del espacio de los polinomios complejos.
3. El conjunto que forman las matrices reales simétricas de tamaño 7×7 es un subespacio del espacio vectorial de las matrices reales de tamaño 7×7.

INTERSECCION DE SUBESPACIOS

[055]

La intersección de subespacios, cualesquiera, de un espacio vectorial V es, a su vez, un subespacio de V.

DEMOSTRACION

Llamemos U a la intersección y sea U_i uno cualquiera de los subespacios que se intersecan. Nótese en primer lugar que el vector nulo $\bar{o}$ pertenece a todos los U_i, luego $\bar{o} \in U$ y entonces es $U \neq \emptyset$. Si $\bar{u}$, $\bar{v} \in U$ y para cualesquiera escalares λ y μ, como $\bar{u}$ y $\bar{v}$ pertenecen a todos los U_i y éstos son subespacios, resulta que $\lambda\bar{u} + \mu\bar{v}$ pertenece a todos los U_i, luego también es de U, como había que comprobar.

EJEMPLO

Sea V el espacio vectorial real de las funciones de $\mathbb{R}$ en $\mathbb{R}$. Sea U_1 el subespacio de V que forman las funciones acotadas y su U_2 el subespacio de V que forman las funciones polinómicas. La intersección $U_1 \cap U_2$ es un subespacio, que está formado por las funciones constantes.

OBSERVACION

La unión de subespacios, de un espacio vectorial V, en general no es un subespacio de V. Así, por ejemplo, en el espacio vectorial $\mathbb{R}^2$, considérense los subespacios U_1 y U_2 siguientes:

$$U_1 = \{(x, 0) \in \mathbb{R}^2 / x \in \mathbb{R}\} \quad \text{y} \quad U_2 = \{(0, y) \in \mathbb{R}^2 / y \in \mathbb{R}\}$$

El conjunto $U_1 \cup U_2$ no es subespacio ya que, por ejemplo, $\vec{u}_1 = (1, 0)$ y $\vec{u}_2 = (0, 1)$ son vectores de $U_1 \cup U_2$ y, sin embargo, su suma $\vec{u}_1 + \vec{u}_2 = (1, 1)$ no pertenece a $U_1 \cup U_2$.

5.3. DEPENDENCIA E INDEPENDENCIA LINEAL

COMBINACIONES LINEALES

[056]

Definición.—Sea V un espacio vectorial sobre un cuerpo K; sea $S = (\vec{u}_1, \vec{u}_2, ..., \vec{u}_p)$(*) un sistema de vectores de V. Se llaman *combinaciones lineales* de los vectores de S a los vectores

$$\vec{v} = \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \cdots + \lambda_p \vec{u}_p$$

donde $\lambda_1, \lambda_2, ..., \lambda_p$ son escalares cualesquiera (del cuerpo K).

Teorema.—El conjunto de todas las combinaciones de los vectores de S es un subespacio vectorial, que se llama *subespacio engendrado* por S y se denota poniendo $\mathcal{V}(S)$ ó

$$\mathcal{V}(\vec{u}_1, \vec{u}_2, ..., \vec{u}_p) = \{\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \cdots + \lambda_p \vec{u}_p / \lambda_i \in K \text{ para } i = 1, 2, ..., p\}^{(**)}$$

También se dice que S es un *sistema generador* de $\mathcal{V}(S)$. Este subespacio $\mathcal{V}(S)$ es el menor de todos los subespacios que incluyen a S.

(*) S es un conjunto de vectores dados en un cierto orden.

(**) Este subespacio, que otros denotan poniendo $<S>$ o $[S]$, también se suele llamar «envolvente lineal de S" «clausura lineal de S».

Si S es un sistema de infinitos vectores de un espacio V, se llama $\mathcal{V}(S)$ al conjunto de todas las combinaciones lineales de cualquier cantidad finita de vectores de V. Se verifica que $\mathcal{V}(S)$ es un subespacio de V; este subespacio $\mathcal{V}(S)$ es el menor subespacio que incluye a S y se llama subespacio engendrado por S.

DEMOSTRACION

- Veamos primero que $\mathcal{V}(S)$ es un subespacio, esto es, que si $\vec{v}, \vec{w} \in \mathcal{V}(S)$ también $\alpha\vec{v} + \beta\vec{w}$ es de $\mathcal{V}(S)$ para cualesquiera $\alpha, \beta \in K$. Así ocurre, ya que, por la definición de $\mathcal{V}(S)$, existen ciertos $\lambda_i, \mu_i \in K$ $(i = 1, 2, ..., p)$ tales que:

$$\vec{v}, \vec{w} \in \mathcal{V}(S) \Rightarrow \vec{v} = \lambda_1\vec{u}_1 + \cdots + \lambda_p\vec{u}_p \quad \text{y} \quad \vec{w} = \mu_1\vec{u}_1 + \cdots + \mu_p\vec{u}_p \Rightarrow$$
$$\Rightarrow \alpha\vec{v} + \beta\vec{w} = (\alpha\lambda_1 + \beta\mu_1)\vec{u}_1 + \cdots + (\alpha\lambda_p + \beta\mu_p)\vec{u}_p \Rightarrow$$
$$\Rightarrow \alpha\vec{v} + \beta\vec{w} \in \mathcal{V}(S)$$

- Todo subespacio que incluya a S ha de contener a las combinaciones lineales de los vectores de S, esto es, ha de incluir a $\mathcal{V}(S)$. Como, además, $\mathcal{V}(S)$ es subespacio que incluye a S, se concluye de todo ello que $\mathcal{V}(S)$ es el menor subespacio que incluye a S.

EJEMPLOS

1. En el espacio vectorial $\mathbb{R}^4$, los vectores $\vec{u} = (1, 2, 0, 0)$, $\vec{v} = (0, 3, -1, 0)$ y $\vec{w} = (0, 0, 5, 4)$ engendran el subespacio:

$$\mathcal{V}(\vec{u}, \vec{v}, \vec{w}) = \{(\lambda, 2\lambda + 3\mu, -\mu + 5\nu, 4\nu)/\lambda, \mu, \nu \in \mathbb{R}\}$$

2. En el espacio vectorial $\mathbb{R}[x]$, de los polinomios reales con una indeterminada x, considérense los polinomios:

$$p(x) = 1 - x \quad ; \quad q(x) = 1 + x^2 \quad \text{y} \quad r(x) = 2x + x^2$$

El subespacio vectorial engendrado por $p(x)$, $q(x)$ y $r(x)$ está formado por todos los polinomios siguientes (al variar $\alpha, \beta, \gamma \in \mathbb{R}$):

$$\alpha p(x) + \beta q(x) + \gamma r(x) = (\alpha + \beta) + (-\alpha + 2\gamma)x + (\beta + \gamma)x^2 \qquad [I]$$

Nótese que dicho subespacio no es otro que el de los polinomios de grado menor o igual que 2, ya que cualquiera de estos polinomios $a + bx + cx^2$ puede expresarse en la forma [I], sin más que tomar $\alpha = 2a + b - 2c$, $\beta = -a - b + 2c$ y $\gamma = a + b - c$.

SISTEMAS EQUIVALENTES DE VECTORES

[057]

Sean $S = (\vec{u}_1, \vec{u}_2, ..., \vec{u}_p)$ y $T = (\vec{v}_1, \vec{v}_2, ..., \vec{v}_q)$ dos sistemas de vectores de un espacio vectorial V. Se dice que S y T son *sistemas equivalentes* si engendran el mismo subespacio, $\mathcal{V}(S) = \mathcal{V}(T)$, esto es, si todo vector de uno cualquiera de los sistemas $S \circ T$ depende linealmente de los vectores del otro. Esta relación (entre sistemas de vectores de V) es evidentemente una equivalencia, es decir, es reflexiva, simétrica y transitiva.

EJEMPLO

Considérense los siguientes vectores de $\mathbb{R}^3$:

$$\vec{a} = (1, 0, 1) \quad ; \quad \vec{b} = (0, 1, 1) \quad ; \quad \vec{c} = (1, 1, 2)$$
$$\vec{u} = (2, 1, 3) \quad \text{y} \quad \vec{v} = (1, 2, 3)$$

Según se comprueba fácilmente los subespacios $\mathcal{V}(\vec{a}, \vec{b}, \vec{c})$ y $\mathcal{V}(\vec{u}, \vec{v})$ son

iguales, ya que ambos están formados por los vectores $(x, y, z) \in \mathbb{R}^3$ tales que $x + y = z$. Por tanto $(\bar{a}, \bar{b}, \bar{c})$ y $(\bar{u}, \bar{v})$ son sistemas equivalentes.

VECTORES DEPENDIENTES E INDEPENDIENTES

Se dice que unos ciertos vectores son linealmente dependientes si entre ellos existe alguna relación de tipo lineal, es decir, si están ligados linealmente o, más precisamente, si alguna combinación lineal de ellos (con coeficientes no nulos) da como resultado el vector nulo. En caso contrario, los vectores se dirán linealmente independientes.

[058]

Sea $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ un sistema de vectores de un espacio vectorial V sobre un cuerpo K.

- Se dice que S es un sistema *linealmente independiente* o sistema *libre* si la única combinación lineal de ellos que vale $\bar{o}$ es la que tiene todos sus coeficientes nulos; esto es, si

$$\left.\begin{array}{c} \lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_p \bar{u}_p = \bar{o} \\ (\lambda_1, \lambda_2, ..., \lambda_p \in K) \end{array}\right\} \Rightarrow \lambda_1 = 0, \lambda_2 = 0, ..., \lambda_p = 0$$

- Se dice que S es un sistema *linealmente dependiente* o sistema *ligado* si no es un sistema libre, esto es, si existen algunos escalares $\lambda_1, \lambda_2, ..., \lambda_p$ no todos nulos tales que $\lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_p \bar{u}_p = \bar{o}$.

Se dice que un vector *depende linealmente* de otros si aquél es igual a una combinación lineal de estos.

Sea S un sistema de infinitos vectores de un espacio vectorial. Se dice que S es un sistema linealmente independiente o libre si son linealmente independientes todos los subconjuntos finitos de S. Si alguno de estos subconjuntos es linealmente dependiente, se dice que S es linealmente dependiente.

EJEMPLOS

1. En $\mathbb{R}^3$, los vectores $\bar{u} = (2, 1, -5)$, $\bar{v} = (1, -4, 2)$ y $\bar{w} = (1, 2, -4)$ forman un sistema linealmente dependiente, ya que

$$2 \cdot (2, 1, -5) - 1 \cdot (1, -4, 2) - 3 \cdot (1, 2, -4) = (0, 0, 0)$$

es decir $\lambda\bar{u} + \mu\bar{v} + \nu\bar{w} = \bar{o}$ para $\lambda = 2$, $\mu = -1$ y $\nu = -3$ que no son todos nulos.

2. En $\mathbb{R}^3$, los vectores $\bar{u} = (1, 0, 0)$, $\bar{v} = (1, 1, 0)$ y $\bar{w} = (1, 1, 1)$ son linealmente independientes. En efecto, la relación $\lambda\bar{u} + \mu\bar{v} + \nu\bar{w} = \bar{o}$ equivale a

$$(\lambda, 0, 0) + (\mu, \mu, 0) + (\nu, \nu, \nu) = (0, 0, 0) \quad \text{o sea} \quad \begin{cases} \lambda + \mu + \nu = 0 \\ \mu + \nu = 0 \\ \nu = 0 \end{cases}$$

cuya única solución es la $\nu = 0$, $\mu = 0$ y $\lambda = 0$.

3. En el espacio vectorial real $\mathcal{F}(\mathbb{R}, \mathbb{R})$, de las funciones de $\mathbb{R}$ en $\mathbb{R}$, el sistema (sen x, cos x, sen $(\pi/4 + x)$) es linealmente dependiente ya que, como

$$\operatorname{sen}\left(\frac{\pi}{4}+x\right)=\operatorname{sen}\frac{\pi}{4}\cos x+\cos\frac{\pi}{4}\operatorname{sen}x=\frac{\sqrt{2}}{2}\cos x+\frac{\sqrt{2}}{2}\operatorname{sen}x$$

resulta que entre las tres funciones existe una combinación lineal de coeficiente no nulo que es igual a la función nula:

$$\sqrt{2}\cos x+\sqrt{2}\operatorname{sen}x-2\operatorname{sen}(\pi/2+x)=0 \quad , \quad \forall x\in\mathbb{R}$$

OBSERVACIONES

- Si uno de los vectores de un sistema es el vector nulo, entonces el sistema es linealmente dependiente.
- Si $\bar{u} \neq \bar{o}$, entonces el sistema $S = (\bar{u})$ es linealmente independiente. Un sistema $(\bar{u}, \bar{v})$, formado por dos vectores, es linealmente dependiente sí y sólo si uno de ellos es proporcional al otro.
- Si un sistema S de vectores es linealmente dependiente, entonces también lo es cualquier sistema que resulte de añadir algún vector a S.
- Si un sistema S de vectores es linealmente independiente, entonces también lo es cualquier sistema que resulte de prescindir de alguno de los vectores de S.

PROPIEDADES DE LA DEPENDENCIA Y DE LA INDEPENDENCIA

[059]

Se consideran aquí vectores de un espacio vectorial V sobre un cuerpo K:

1. Un sistema de vectores $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es linealmente dependiente sí y sólo si alguno de sus vectores depende linealmente de los demás.
2. Si un vector $\bar{w}$ depende linealmente de los vectores $\bar{u}_1, \bar{u}_2, ..., \bar{u}_p$ y éstos dependen, a su vez, de otros vectores $\bar{v}_1, \bar{v}_2, ..., \bar{v}_q$, entonces $\bar{w}$ depende de éstos últimos.
3. Dos sistemas de vectores $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ y $T = (\bar{v}_1, \bar{v}_2, ..., \bar{v}_p)$ son equivalentes, o sea, $\mathcal{V}(S) = \mathcal{V}(T)$, sí y sólo si todo vector de uno de los sistemas depende linealmente de los vectores del otro, y recíprocamente.
4. Si S es un sistema linealmente independiente de vectores y $\bar{v}$ es un vector que no depende de los vectores de S, entonces el sistema $S \cup \{\bar{v}\}$ también es linealmente independiente.

DEMOSTRACION

1. Si el sistema S es linealmente dependiente, entonces existen algunos escalares $\lambda_1, \lambda_2, ..., \lambda_p$ no todos nulos ($\lambda_i \neq 0$ para algún i) tales que:

$$\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p = \bar{o}$$

Podemos suponer que es $\lambda_1 = 0$ (intercambiando los subíndices 1 e i, si fuera preciso), lo que nos permite poner:

$$\bar{u}_1 = \left(-\frac{\lambda_2}{\lambda_1}\right)\bar{u}_2 + \cdots + \left(-\frac{\lambda_p}{\lambda_1}\right)\bar{u}_p$$

luego uno de los vectores de S, el $\bar{u}_1$, depende linealmente de los demás. Recíprocamente, si uno de los vectores de S. que podemos suponer que es $\bar{u}_1$, depende linealmente de los demás, resulta que (para ciertos escalares $\mu_2, ..., \mu_p$) ha de ser:

$$\bar{u}_1 = \mu_2\bar{u}_2 + \cdots + \mu_p\bar{u}_p$$

luego,

$$(-1)\bar{u}_1 + \mu_2\bar{u}_2 + \cdots + \mu_p\bar{u}_p = \bar{o}$$

Como este último primer coeficiente es $-1 \neq 0$, resulta que los vectores de S son linealmente dependientes.

2. Se sabe que para ciertos escalares $\lambda_1, \lambda_2, ... \lambda_p$ se verifica que

$$\bar{w} = \lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p = \sum_{i=1}^{p} \lambda_i\bar{u}_i$$

También se sabe que, para cada $i = 1, 2, ..., p$, existen ciertos escalares μ_{i1}, $\mu_{i2}, ..., \mu_{iq}$ tales que:

$$\bar{u}_i = \mu_{i2}\bar{v}_1 + \mu_{i2}\bar{v}_2 + \cdots + \mu_{iq}\bar{v}_q = \sum_{j=1}^{q} \mu_{ij}\bar{v}_j$$

De ello se deduce entonces que:

$$\bar{u} = \sum_{i=1}^{p} \lambda_i\bar{u}_i = \sum_{j=1}^{p} \lambda_i\left(\sum_{j=1}^{q} \mu_{ij}\bar{v}_j\right) = \sum_{i=1}^{p}\sum_{j=1}^{q} \lambda_i\mu_{ij}\bar{v}_j = \sum_{j=1}^{q}\left(\sum_{i=1}^{p} \lambda_i\mu_{ij}\right)\bar{v}_j$$

es decir,

$$\bar{u} = \sum_{j=1}^{q} \alpha_j\bar{v}_j$$

donde

$$\alpha_j = \sum_{i=1}^{p} \lambda_i \mu_{ij}$$

luego $\bar{u}$ depende linealmente de los $\bar{v}_1, \bar{v}_2, ..., \bar{v}_q$.

3. Si $\mathcal{V}(S) = \mathcal{V}(T)$, entonces todo vector de cualquiera de estos dos subespacios, en particular los vectores que lo engendran, pertenecen también al otro subespacio, o sea, dependen linealmente de los vectores que engendran a este último, que es lo que se deseaba comprobar. Recíprocamente, si los vectores de S dependen linealmente de los de T, como todo vector de $\mathcal{V}(S)$ depende linealmente de los vectores de S, según la propiedad anterior resulta que los vectores de $\mathcal{V}(S)$ depende linealmente de los vectores de T, es decir, que $\mathcal{V}(S) \subset \mathcal{V}(T)$. Como, además, se sabe que los vectores de T dependen linealmente de los de T, nos encontramos con que también se verifica que $\mathcal{V}(T) \subset \mathcal{V}(S)$. Por tanto, es $\mathcal{V}(S) = \mathcal{V}(T)$, como había de probar.
4. Sea $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$. Suponiendo que

$$\lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_p \bar{u}_p + \mu \bar{v} = \bar{o}$$

hay que probar que es $\mu = \lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$. Si fuese $\mu \neq 0$, resultaría que

$$\bar{v} = \left(-\frac{\lambda_1}{\mu}\right)\bar{u}_1 + \left(-\frac{\lambda_2}{\mu}\right)\bar{u}_2 + \cdots + \left(-\frac{\lambda_p}{\mu}\right)\bar{u}_p$$

que es falso, pues $\bar{v}$ no depende linealmente de los $\bar{u}_1, \bar{u}_2, ..., \bar{u}_p$; por tanto, concluimos que ha de ser $\mu = 0$ y, en consecuencia, se verifica que

$$\lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_p \bar{u}_p = \bar{o}$$

Ahora bien, como $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es un sistema linealmente independiente, de esta última relación se desprende que $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$, con lo que concluye la demostración.

TEOREMA FUNDAMENTAL DE LA INDEPENDENCIA LINEAL

[060]

Sea V un espacio vectorial que está engendrado por un cierto sistema $G = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$, de un número finito p de vectores. Si $I = (\bar{v}_1, \bar{v}_2, ..., \bar{v}_h)$ es un sistema independiente, formado por h vectores de V, entonces se verifica que $h \leq p$.

DEMOSTRACION

Vamos a proceder por reducción al absurdo: suponiendo que es $p < h$ obtendremos una contradicción.

Por ser G un sistema generador de V, el vector $\bar{v}_1 \in V$ se puede expresar en la forma $\bar{v}_1 = \lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p$, donde alguno de los escalares λ_i es no nulo ya que $\bar{v}_1 \neq \bar{o}$ (pues $\bar{v}_1 \in I$ e I es independiente) y podemos suponer, pues, que es $\lambda_1 \neq 0$; por tanto, de la anterior igualdad se puede despejar $\bar{u}_1$, que resulta ser, así, combinación lineal de los $\bar{v}_1, \bar{u}_2, \bar{u}_3, \ldots, \bar{u}_p$. Como $G = (\bar{u}_1, \bar{u}_2, \bar{u}_3, \ldots, \bar{u}_p)$ es generador de V, de lo dicho se desprende que:

$$G_1 = (\bar{v}_1, \bar{u}_2, \bar{u}_3, \ldots, \bar{u}_p)$$

es generador de V. Por ser G_1 un sistema generador de V, el vector $\bar{v}_2 \in V$ se puede expresar en la forma $\bar{v}_2 = \mu_1\bar{v}_1 + \mu_2\bar{u}_2 + \mu_3\bar{u}_3 + \cdots + \mu_p\bar{u}_p$, donde los escalares $\mu_2, \mu_3, \ldots, \mu_p$ no pueden ser todos nulos, pues $\bar{v}_1$ y $\bar{v}_2$ son independientes (pues ambos son de I) y podemos suponer, pues, que es $\mu_2 \neq 0$; por tanto, de la anterior igualdad se puede despejar $\bar{u}_2$, que resulta ser, así, combinación lineal de los $\bar{v}_1, \bar{v}_2, \bar{u}_3, \ldots, \bar{u}_p$. Como $G_1 = (\bar{v}_1, \bar{u}_2, \bar{u}_3, \ldots, \bar{u}_p)$ es generador de V, de lo dicho se desprende que:

$$G_2 = (\bar{v}_1, \bar{v}_2, \bar{u}_3, \ldots, \bar{u}_p)$$

es generador de V. Reiterando este proceso p veces, se obtiene finalmente que:

$$G_p = (\bar{v}_1, \bar{v}_2, \bar{v}_3, \ldots, \bar{v}_p)$$

es generador de V. Por tanto, el vector $\bar{v}_{p+1}$, que hemos supuesto que existe al admitir que es $h > p$, ha de ser una combinación lineal de los vectores de G_p, lo cual no es posible debido a que $I = (\bar{v}_1, \bar{v}_2, \bar{v}_3, \ldots, \bar{v}_h)$ es un sistema independiente. Esta contradicción obliga a desechar la relación $p < h$.

β BASES. COORDENADAS

5.4. ESPACIOS DE DIMENSION FINITA

Nos vamos a ocupar aquí de aquellos espacios vectoriales que están generados por un número finito de vectores, esto es, tales que todos sus vectores son combinaciones lineales de un número finito de ellos. Estos espacios se llamarán de dimensión finita aunque, de momento y hasta que hayamos definido lo que es la dimensión, los llamaremos de tipo finito.

En un espacio vectorial de tipo finito, tienen especial interés aquellos sistemas generadores que son, además, independientes; a ellos se los llama bases. Todo vector del espacio se podrá poner de una única manera como combinación lineal de los vectores de una base; a los coeficientes de esta combinación se les llama coordenadas del vector dado.

ESPACIOS DE TIPO FINITO. BASES

[061]

Un espacio vectorial V, sobre un cuerpo K, se dice que es de *tipo finito* si está generado por un número finito de vectores, es decir, si en V existe algún sistema de vectores $S = (u_1, u_2, ..., u_p)$ tal que $V = \mathcal{V}(S)$.

Si V es de tipo finito, se dice que un sistema de vectores $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una *base* de V si se verifica una cualquiera de las dos condiciones siguientes, que son equivalentes entre sí (como después se prueba):

1.ª Definición: B es un sistema generador de V que, además, es sistema linealmente independiente.

2.ª Definición: todo vector de V se puede expresar de una sola manera como combinación lineal de los vectores de B.

DEMOSTRACION

a) Comprobaremos primeramente que 1.ª $\Rightarrow$ 2.ª. Como B es sistema generador de V, todo vector $\bar{v} \in V$ es combinación lineal de los vectores de B; veamos que esta combinación es única. Supongamos que hubiera más de una:

$$\bar{v} = \lambda_1\bar{e}_1 + \lambda_2\bar{e}_2 + \cdots + \lambda_n\bar{e}_n \quad \text{y} \quad \bar{v} = \mu_1\bar{e}_1 + \mu_2\bar{e}_2 + \cdots + \mu_n\bar{e}_n$$

Restando miembro a miembro, se obtendría que:

$$(\lambda_1 - \mu_1)\bar{e}_1 + (\lambda_2 - \mu_2)\bar{e}_2 + \cdots + (\lambda_n - \mu_n)\bar{e}_n = \bar{o}$$

Como se ha supuesto que B es un sistema independiente, resulta que los coeficientes de esta última combinación lineal son todos nulos, es decir, será:

$$\lambda_1 = \mu_1,\ \lambda_2 = \mu_2,\ ...,\ \lambda_n = \mu_n$$

Luego las dos expresiones $\bar{v} = \Sigma\,\lambda_i\bar{e}_i$ y $\bar{v} = \Sigma\,\mu_i\bar{e}_i$ son la misma, como había que comprobar.

Se dice que un sistema S de infinitos vectores de un espacio V es una *base infinita* de este si S es independiente y genera todo V. Esto es si: 1) cualquier subconjunto finito de S es linealmente independiente, y 2) cualquier vector de V es una combinación lineal de algún subconjunto finito de S.

b) Veamos ahora que 2.ª $\Rightarrow$ 1.ª. Como todo $\bar{v} \in V$ es combinación lineal de los vectores de B, este sistema B es generaodr de V. Comprobemos que B es, también, independiente. Supongamos, para ello, que existieran ciertos escalares $\lambda_1, \lambda_2, ..., \lambda_n$ tales que

$$\bar{o} = \lambda_1\bar{e}_1 + \lambda_2\bar{e}_2 + \cdots + \lambda_n\bar{e}_n \qquad [I]$$

Si recurrimos a que, obvimente, se verifica que:

$$\bar{o} = 0\bar{e}_1 + 0\bar{e}_2 + \cdots + 0\bar{e}_n \qquad [II]$$

nos encontramos con que [I] y [II] son dos maneras de expresar el vector $\bar{o}$ como combinación lineal de los vectores de B; por tanto, según nuestra hipótesis ambas coinciden, es decir, se verifica que

$$\lambda_1 = 0,\ \lambda_2 = 0,\ ...,\ \lambda_n = 0$$

lo que permite asegurar que B es un sistema linealmente independiente.

EJEMPLOS

1. En el espacio vectorial K^n (en el que K es el cuerpo de escalares), los n vectores:

$$\bar{e}_1 = (1, 0, 0, ..., 0),\ \bar{e}_2 = (0, 1, 0, ..., 0),\ ...,\ \bar{e}_n = (0, 0, 0, ..., 1)$$

forman una base, que se llama *base canónica* de K^n. Así ocurre, en efecto, ya que $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es:

- Sistema generador, puesto que cualquier vector $(x_1, x_2, ..., x_n)$ de K^n es combinación lineal de los vectores de B, ya que:

$$(x_1, x_2, ..., x_n) = x_1(1, 0, ..., 0) + x_2(0, 1, ..., 0) + \cdots + x_n(0, 0, ..., 1)$$

- Sistema independiente, ya que la relación $\lambda_1\bar{e}_1 + \lambda_2\bar{e}_2 + \cdots + \lambda_n\bar{e}_n = \bar{o}$ equivale a $(\lambda_1, \lambda_2, ..., \lambda_n) = (0, 0, ..., 0)$; es decir, a $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

2. En el espacio vectorial $\mathcal{M}_{m\times n}$, de las matrices reales de tamaño $m \times n$, las $m \cdot n$ siguientes matrices E_{ij} (para $i = 1, 2, ..., m$ y $j = 1, 2, ..., n$) forman una base, siendo:

E_{ij} = matriz que tiene nulos todos sus elementos excepto el que ocupa el lugar ij, que vale la unidad.

En efecto: $B = \{E_{ij}/i \in 1, 2, ..., m;\ j = 1, 2, ..., n\}$ es generador e independiente, pues:

- Cualquier matriz $A = [a_{ij}]$ de $\mathcal{M}_{m\times n}$ es combinación lineal de las matrices de B, ya que

$$A = [a_{ij}] = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij}$$

- La única combinación lineal nula de las matrices E_{ij} es la que tiene todos los coeficientes nulos, ya que:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_{ij}E_{ij} = O \Rightarrow [\lambda_{ij}] = O \Rightarrow$$
$$\Rightarrow \lambda_{ij} = 0 \quad (i = 1, 2, ..., m;\ j = 1, 2, ..., n)$$

A esta base $\{E_{ij}/i = 1, 2, ..., m; j = 1, 2, ..., n\}$ la llamaremos *base usual* de $\mathcal{M}_{m\times n}$.

3. En el espacio vectorial V de los polinomios reales de grado menor o igual que n (con una indeterminada x), el siguiente sistema B de $n+1$ polinomios es una base:

$$B = (1, x, x^2, ..., x^n)$$

Así ocurre ya que B es independiente y genera todo V, como se comprueba trivialmente. A esta base B la llamaremos *base usual* del espacio vectorial de los polinomios de grado menor o igual que n.

EXISTENCIA DE BASES

[062]

Cualquier sistema generador de un espacio vectorial de tipo finito, $V \neq O$, incluye a una base de V. En consecuencia, todo espacio vectorial $V \neq O$ de tipo finito tiene alguna base.

DEMOSTRACION

Sea $G = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ cualquier sistema finito de vectores que engendre a V; un tal sistema G existe ya que V es un espacio de tipo finito. Si los vectores de G son linealmente independientes, entonces G es base de V y el teorema es cierto. De no ser así, alguno de los vectores de G depende linealmente de los demás; si es $\bar{u}_i$ dicho vector, el sistema $G^{(1)} = G - \{\bar{u}_i\}$ engendra también al espacio V. Si $G^{(1)}$ fuese linealmente independiente, entonces sería una base de V y el teorema se cumpliría. Si $G^{(1)}$ fuese linealmente dependiente, procediendo como ya se hizo con G, se obtiene un nuevo sistema $G^{(2)} = G^{(1)} - \{\bar{u}_j\}$ que es generador de V. Al reiterar este proceso, obteniendo así sucesivos sistemas G, $G^{(1)}$, $G^{(2)}$, $G^{(3)}$, ..., llega un momento en el que se obtiene una base de V y ello ocurre antes de llegar a un sistema $G^{(p)}$, pues si se llegase (en el peor de los casos) a un $G^{(p-1)}$, como este sistema tendría un solo vector, y generaría todo $V \neq O$, dicho vector no podría ser nulo, luego $G^{(p-1)}$ sería base de V.

TEOREMA DE LA DIMENSION

[063]

Todas las bases de un espacio vectorial $V \neq O$, de tipo finito, tienen igual número de vectores. A este número se le llama *dimensión* del espacio V y se le representa poniendo dim V (*).

Se conviene en que el espacio $V = O$ tiene dimensión 0.

(*) Una vez que hemos llegado a definir lo que es la dimensión en los espacios de tipo finito, a estos espacios los llamaremos en adelante «de dimensión finita».

DEMOSTRACION

Hay que probar que para cualesquiera que sean las bases B y B' de V, si n es el número de elementos de B y n' es el número de elementos de B', se verifica que $n = n'$. En efecto:

Como B es un sistema generador de V y B' es un sistema independiente, del teorema fundamental de la independencia lineal se desprende que $n' \leqslant n$. Intercambiando ahora los papeles de B y B', se obtiene análogamente que $n \leqslant n'$. Por tanto, ha de ser $n = n'$.

PROPIEDADES DE LAS BASES Y LA DIMENSION

[064]

Si $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es un sistema de vectores de un espacio vectorial V de dimensión finita, entonces se verifica que:

1. **Si $S = (\bar{u}_1, ..., \bar{u}_p)$ es un sistema generador de V, entonces $p \geqslant \dim V$.**
2. **Si $S = (\bar{u}_1, ..., \bar{u}_p)$ es un sistema independiente, entonces $p \leqslant \dim V$.**
3. **Si $S = (\bar{u}_1, ..., \bar{u}_p)$ es generador de V y $\dim V = p$, entonces S es base del espacio V.**
4. **Si $S = (\bar{u}_1, ..., \bar{u}_p)$ es independiente y $\dim V = p$, entonces S es base del espacio V.**

DEMOSTRACION

1. Sabemos que S, por ser un sistema generador de V, incluye a una base de V y, por ello, la dimensión de V, que es el número de vectores de una base, no puede ser mayor que el número p de vectores de S, es decir, ha de ser $\dim V \leqslant p$.
2. Recurramos a que una base B de V es un sistema generador. Según el teorema fundamental de la independencia lineal (véase [060]), como S es independiente y B es generador, el número p de vectores de S es menor o igual que el número de vectores de B, o sea, ha de ser $p \leqslant \dim V$.
3. Como S es un sistema generador de V, sabemos que S incluye a una base de V; una tal base debe de ser todo S, ya que el número de vectores de cualquier base es $\dim V$, que es igual al número p de vectores que tiene S.
4. Supongamos que S no es base, es decir, que no genera todo V. Habría entonces un vector $\bar{u}_{p+1}$ independiente de S y, por ello, $S' = S \cup \{\bar{u}_{p+1}\}$ sería un sistema independiente. De ser así y de acuerdo con la propiedad 2, como S' es independiente y tiene $p + 1$ vectores, se verificaría que $p + 1 \leqslant \dim V$, que es falso dado que $p = \dim V$. Esta contradicción prueba que S es base de V.

OBSERVACIONES

1. La dimensión de un espacio vectorial V es el número máximo de vectores de V linealmente independientes.

2. La dimensión de un espacio vectorial V es el número mínimo de vectores de un sistema generador de V.

TEOREMA DE LA BASE INCOMPLETA

[065]

En un espacio vectorial V de dimensión finita, todo sistema linealmente independiente de vectores puede completarse hasta obtener una base. Dicho con mayor precisión:

Si V es un espacio vectorial de dimensión finita n y si $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es un sistema independiente de p vectores de V, donde $p < n$, entonces es posible encontrar (mejor dicho, existe) algún sistema S' de $n - p$ vectores de V, tal que $S \cup S'$ sea una base de V. Es más, los vectores de S' se pueden tomar de entre los de una base dada cualquiera $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V.

DEMOSTRACION

Obviamente, esta proposición se habrá comprobado en su totalidad si se prueba que, dada la base $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$, los vectores de S' se pueden elegir de entre los de B. Veamos si esto es cierto.

En B hay algún vector $\bar{e}_i$ que no depende linealmente de S, pues de lo contrario, como B es generador de V, también lo sería S y esto es falso ya que $p < n$. El sistema $S \cup \{\bar{e}_i\}$ de $p + 1$ vectores es, pues, linealmente independiente. Si $p + 1 = n$, este sistema $S \cup \{\bar{e}_i\}$ es base y la propiedad queda probada; si $p + 1 < n$, el proceso continúa.

En $B - \{\bar{e}_i\}$ hay algún vector $\bar{e}_j$ que no depende linealmente de $S \cup \{\bar{e}_i\}$ pues de lo contrario, como B es generador de V, también lo sería $S \cup \{\bar{e}_i\}$ y esto es falso ya que $p + 1 < n$. El sistema $S \cup \{\bar{e}_i, \bar{e}_j\}$ de $p + 2$ vectores es, pues, linealmente independiente. Si $p + 2 = n$, este sistema $S \cup \{\bar{e}_i, \bar{e}_j\}$ es base y la propiedad queda probada; si $p + 2 < n$, el proceso continúa.

Al cabo de $n - p$ etapas, este camino desemboca en el sistema $S \cup S'$ que se buscaba.

EJEMPLO

En el espacio vectorial $\mathbb{R}^3$, los vectores $\bar{a} = (1, -1, 2)$ y $\bar{b} = (0, 1, -2)$ son linealmente independientes y, por ello, es seguro que hay algún vector $\bar{c} \in \mathbb{R}^3$ tal que $(\bar{a}, \bar{b}, \bar{c})$ es una base. Es más, el vector $\bar{c}$, que se puede elegir de muchas maneras, puede ser, en particular, uno de los tres vectores de la base canónica (o también de cualquier otra base), que son los $\bar{e}_1 = (1, 0, 0)$, $\bar{e}_2 = (0, 1, 0)$ y $\bar{e}_3 = (0, 0, 1)$. Nótese que $\bar{c}$ no puede ser $\bar{e}_1$, ya que $\bar{e}_1$ es combinación lineal de $\bar{a}$ y $\bar{b}$, pues $\bar{e}_1 = \bar{a} + \bar{b}$. Sin embargo, sí puede tomarse $\bar{c} = \bar{e}_2$ pues, según resulta fácil comprobar, $\bar{e}_2$ es linealmente independiente de los $\bar{a}$ y $\bar{b}$.

DIMENSION DE UN SUBESPACIO

[066]

Si U es un subespacio de un espacio vectorial V y si V tiene dimensión finita, entonces U también tiene dimensión finita y se verifica que $\dim U \leq \dim V$, donde el signo $=$ sólo es válido en el caso de ser $U = V$.

DEMOSTRACION

Si $U = O$ la propiedad es evidente; se supondrá, pues, que U tiene algún vector no nulo. En U hay sistemas independientes y todos ellos tienen no más de $n = \dim V$ vectores, ya que dichos sistemas lo son de vectores independientes de V. Si m es el número máximo de vectores que resultan tener tales sistemas, es $m = \dim U$. Como, según se ha dicho, es $m \leq n$, se concluye que $\dim U \leq \dim V$. Para que sea $\dim U = \dim V$, esto es, $m = \dim V$, es necesario y suficiente que U incluya a un sistema independiente (de m vectores) que sea base de V, o sea, que $U = V$.

5.5. COORDENADAS

Lo que hace del concepto de base algo realmente útil es que, recurriendo a una de ellas, cualquier vector queda identificado mediante los coeficientes de la única combinación lineal que lo expresa en función de los vectores de aquélla. A estos coeficientes se les llama coordenadas. En un espacio de dimensión finita, si se dispone de una base, conocer un vector viene a ser lo mismo que conocer sus coordenadas.

[067]

Sea V un espacio vectorial de dimensión finita, sobre un cuerpo K. Dada una base $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V, se sabe que para cada vector $\bar{x} \in V$ existen unos únicos escalares $x_1, x_2, ..., x_n \in K$ tales que:

$$\bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n \qquad (I)$$

Pues bien, se dice que la n-upla o sucesión de n escalares $(x_1, x_2, ..., x_n)$ es el sistema de *coordenadas* del vector $\bar{x}$ en la base B. Cuando se sobreentiende cuál es la base B que se considera, para señalar que $x_1, x_2, ..., x_n$ son las coordenadas(*) de $\bar{x}$, en lugar de recurrir a la expresión (I), se suele poner:

$$\bar{x}(x_1, x_2, ..., x_n) \quad \text{o abreviadamente} \quad \bar{x}(x_i)$$

(*) Se llaman «componentes» del vector $\bar{x}$ respecto de la base B a los n vectores $x_1\bar{e}_1$, $x_2\bar{e}_2$, ..., $x_n\bar{e}_n$.

EJEMPLO

Los vectores $\bar{u}_1 = (1, 0, 0)$, $\bar{u}_2 = (1, 1, 0)$ y $\bar{u}_3 = (1, 1, 1)$ forman una base de $\mathbb{R}^3$, según se comprueba fácilmente. En esta base, las coordenadas de un vector $\bar{v} = (x, y, z)$ son los escalares $x - y$, $y - z$, z, ya que:

$$(x-y)\bar{u}_1 + (y-z)\bar{u}_2 + z\bar{u}_3 = [x-y]\cdot(1,0,0) + [y-z]\cdot(1,1,0) + z\cdot(1,1,1) =$$
$$= (x-y+y-z+z,\ y-z+z,\ z) = (x, y, z) = (x, y, z) = \bar{v}$$

ISOMORFISMO DE COORDENADAS

Vamos a estudiar con algún detalle la aplicación biyectiva que a cada vector $\bar{x}$, de un espacio V de dimensión finita n, le asocia la n-upla $(x_1, x_2, ..., x_n)$ de sus coordenadas, que es un vector del espacio vectorial K^n (donde K es el cuerpo de escalares).

Esta aplicación biyectiva es especialmente interesante ya que permite «identificar» cada vector $\bar{x} \in V$ con su sistema de coordenadas $(x_1, x_2, ..., x_n) \in K^n$. Esto de identificar significa que, prescindiendo de la naturaleza concreta de los vectores de V y salvo cuestiones de notación, no hay nada que diferencie a los vectores de V de los de K^n, de modo que, desde el punto de vista del álgebra, $\bar{x}$ y $(x_1, x_2, ..., x_n)$ pueden considerarse como una misma cosa. Esto último significa que si se opera (suma y producto por escalar) con unos ciertos vectores de V y se hace luego lo mismo con sus sistemas de coordenadas, el vector que se obtiene al principio tiene por coordenadas al resultado de la segunda operación. Para expresar este hecho, se dirá que V y K^n son espacios vectoriales isomorfos. Precisemos ya lo hasta aquí dicho:

[068]

Sea V un espacio vectorial de dimensión n, sobre un cuerpo K. Dada una base $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V, se considera la aplicación «COOR», que a cada vector le asocia su sistema de coordenadas; esto es:

$$V \xrightarrow{\text{COOR}} K^n$$
$$\bar{x} \longmapsto \text{COOR } \bar{x} = (x_1, x_2, ..., x_n)$$

(x_i es la coordenada de lugar i (para $i = 1, 2, ..., n$) de $\bar{x}$ en la base dada B). Esta aplicación es «lineal» para la suma y para el producto por un escalar, esto es, tal que, para cualesquiera $\bar{x}, \bar{y} \in V$ y $\lambda, \mu \in K$, se verifica:

$$\text{COOR } (\lambda\bar{x} + \mu\bar{y}) = \lambda\ (\text{COOR } \bar{x}) + \mu\ (\text{COOR } \bar{y})$$

Hay que notar que la aplicación recíproca de la COOR también es biyectiva y lineal. Por todo ello, se dice que la aplicación COOR es un *isomorfismo* (biyectiva y lineal) del espacio V en el K^n. También se dice que V y K^n son espacios *isomorfos*(*) o identificables.

(*) De iso, igual, y morfo, forma.

DEMOSTRACION

De acuerdo con la definición de la aplicación COOR, se tiene:

$$\text{si} \quad \text{COOR } \bar{x} = (x_1, x_2, ..., x_n), \quad \text{entonces} \quad \bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n = \sum x_i\bar{e}_i$$
$$\text{si} \quad \text{COOR } \bar{y} = (y_1, y_2, ..., y_n), \quad \text{entonces} \quad \bar{y} = y_1\bar{e}_1 + y_2\bar{e}_2 + \cdots + y_n\bar{e}_n = \sum y_i\bar{e}_i$$

De ello resulta que el vector $\lambda\bar{x} + \mu\bar{y}$ se expresará en la forma:

$$\lambda\bar{x} + \mu\bar{y} = \lambda\left(\sum_{i=1}^{n} x_i\bar{e}_i\right) + \mu\left(\sum_{i=1}^{n} y_i\bar{e}_i\right) = \sum_{i=1}^{n} (\lambda x_i + \mu y_i)\bar{e}_i$$

luego la coordenada i-ésima de $\lambda\bar{x} + \mu\bar{y}$ es $\lambda x_i + \mu y_i$, con lo que se obtiene:

$$\text{COOR } (\lambda\bar{x} + \mu\bar{y}) = (\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, ..., \lambda x_n + \mu y_n) =$$
$$= \lambda(x_1, x_2, ..., x_n) + \mu(y_1, y_2, ..., y_n) = \lambda \text{ (COOR } \bar{x}) + \mu \text{ (COOR } \bar{y})$$

Para acabar, comprobemos que la aplicación recíproca de COOR que denotaremos por $f: K^n \to V$, también es lineal, esto es, que:

$$f(\lambda\bar{x}^* + \mu\bar{y}^*) = \lambda f(\bar{x}^*) + \mu f(\bar{y}^*)$$

para cualesquiera $\bar{x}^*, \bar{y}^* \in K^n$ y $\lambda, \mu \in K$. Como COOR es biyectiva, la anterior igualdad es equivalente a la

$$\text{COOR } [f(\lambda\bar{x}^* + \mu\bar{y}^*)] = \text{COOR } [\lambda f(\bar{x}^*) + \mu f(\bar{y}^*)]$$

Teniendo en cuenta que la composición de COOR y f es la identidad y que COOR es lineal, la anterior igualdad equivale a la:

$$\lambda\bar{x}^* + \mu y^* = \lambda \text{ COOR } [f(\bar{x}^*)] + \mu \text{ COOR } [f(\bar{y}^*)] = \lambda\bar{x}^* + \mu\bar{y}^*$$

que es evidentemente cierta.

CAMBIO DE COORDENADAS

Si se dispone de dos bases B y B' en un mismo espacio vectorial V, cada vector $\bar{x} \in V$ tendrá dos sistemas de coordenadas, uno respecto de cada base. Si se conoce una base respecto de la otra, va a ser posible relacionar unas coordenadas con otras; las relaciones que ligan a los dos sistemas de coordenadas se llaman «ecuaciones del cambio de coordenadas». Así, por ejemplo, si $(\bar{e}_1, \bar{e}_2)$ es una base de $\mathbb{R}^2$, el vector $\bar{u} = x_1\bar{e}_1 + x_2\bar{e}_2$, al expresarle en la base que forman los vectores $\bar{e}_1' = \bar{e}_1 + \bar{e}_2$ y $e_2' = \bar{e}_1 - \bar{e}_2$, queda:

$$\bar{u} = x_1 \tfrac{1}{2}(\bar{e}_1' + \bar{e}_2') + x_2 \tfrac{1}{2}(\bar{e}_1' - \bar{e}_2') = \frac{x_1 + x_2}{2}\bar{e}_1' + \frac{x_1 - x_2}{2}\bar{e}_2'$$

esto es, las nuevas coordenadas de $\bar{u}$ son $\frac{1}{2}(x_1 + x_2)$ y $\frac{1}{2}(x_1 - x_2)$.

[069]

Sea V un espacio vectorial de dimensión n, sobre un cuerpo K. Sea $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ una base de V, en la que el sistema de coordenadas de un vector $\bar{x} \in V$ se denota por $(x_1, x_2, ..., x_n)$. Sea $B' = (\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n)$ una nueva base de V, en la que el sistema de coordenadas de $\bar{x}$ es $(x'_1, x'_2, ..., x'_n)$. Si se conocen los vectores de B' en función de los de B, es decir, si se conocen los escalares q_{ij} que permiten poner:

$$\bar{e}'_j = \sum_{i=1}^{n} q_{ij}\bar{e}_i \quad (\text{para } j = 1, 2, ..., n)$$

o sea:

$$\begin{cases} \bar{e}'_1 = q_{11}\bar{e}_1 + q_{21}\bar{e}_2 + \cdots + q_{n1}\bar{e}_n \\ \bar{e}'_2 = q_{12}\bar{e}_1 + q_{22}\bar{e}_2 + \cdots + q_{n2}\bar{e}_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \bar{e}'_n = q_{1n}\bar{e}_1 + q_{2n}\bar{e}_2 + \cdots + q_{nn}\bar{e}_n \end{cases} \quad \text{(I)}$$

entonces el sistema de coordenadas $(x_1, x_2, ..., x_n)$ valdrá, en función del $(x'_1, x'_2, ..., x'_n)$:

$$x_i = \sum_{j=1}^{n} q_{ij}x'_j \quad (\text{para } i = 1, 2, ..., n)$$

o sea:

$$\begin{cases} x_1 = q_{11}x'_1 + q_{12}x'_2 + \cdots + q_{1n}x'_n \\ x_2 = q_{21}x'_1 + q_{22}x'_2 + \cdots + q_{2n}x'_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = q_{n1}x'_1 + q_{n2}x'_2 + \cdots + q_{nn}x'_n \end{cases} \quad \text{(II)}$$

Estas ecuaciones (II) pueden expresarse matricialmente poniendo:

$$X = QX', \quad \text{donde} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad X' = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix} \quad Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix}$$

Q se llama «matriz del cambio de coordenadas»(*) y es una matriz regular o invertible; su matriz inversa es la matriz del cambio de coordenadas inverso, esto es, del que se obtiene al pasar de la base B' a la base B.

(*) La matriz Q es la que gobierna en las ecuaciones (II), que matricialmente son $X = QX'$. En la relación (I) gobierna la matriz Q', traspuesta de Q; a la matriz Q' se la suele llamar matriz del cambio de base. Nótese que no sería correcto expresar las relaciones (I) de forma matricial, pues para poderlo hacer habría que recurrir a algo que no se define: matrices columna cuyos elementos fueran los vectores $\bar{e}_i$ y $\bar{e}'_j$, cosa ésta que carece de sentido, ya que los elementos de cualquier matriz son escalares y no vectores.

DEMOSTRACION

El vector genérico $\bar{x} \in V$ se expresa, en una y otra base, en la forma:

$$\bar{x} = \sum_{i=1}^{n} x_i \bar{e}_i \quad \text{y} \quad \bar{x} = \sum_{j=1}^{n} x'_j \bar{e}'_j \qquad (A)$$

De la segunda de estas relaciones y de las ecuaciones (I), se obtiene:

$$\bar{x} = \sum_{j=1}^{n} x'_j \bar{e}'_j = \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} q_{ij} \bar{e}_i \right) = \sum_{i,j=1}^{n} x'_j q_{ij} \bar{e}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} x'_j q_{ij} \right) \bar{e}_i$$

Esta relación informa de que $\sum_j x'_j q_{ij}$ es la coordenada de lugar i de $\bar{x}$ en la base B, o sea, que es igual a x_i (véase la primera de las dos relaciones (A)); luego

$$x_i = \sum_{j=1}^{n} q_{ij} x'_j \quad (\text{para } i = 1.\ 2,\ ...,\ n)$$

Supongamos ahora que cambiamos nuevamente de base, pasando de la base B' a la base B de partida; llamando P a la matriz de este cambio de coordenadas, se verifica que $X' = PX$. Combinando esta relación con la anterior $X = QX'$, se obtiene que:

$$X = QX' = (QP)X \quad \text{y} \quad X' = PX = (PQ)X'$$

Como estas relaciones son ciertas para cualesquiera matrices columna X y X', resulta que $QP = I$ y $PQ = I$, es decir P y Q son la una inversa de la otra y, consecuentemente, son regulares.

EJERCICIO

Hallar las coordenadas del vector $\bar{u} = (x, y, z) \in \mathbb{R}^3$ en la base $B' = (\bar{e}'_1, \bar{e}'_2, \bar{e}'_3)$, donde

$$\bar{e}'_1 = (1, 2, 0) \quad ; \quad \bar{e}'_2 = (-3, -7, 1) \quad ; \quad \bar{e}'_3 = (0, -2, 1)$$

RESOLUCION

Como x, y, z son las coordenadas de $\bar{u}$ en la base canónica de $\mathbb{R}^3$, llamando x_1, x_2, x_3 a las coordenadas pedidas, resulta que, por ser

$$\begin{cases} \bar{e}'_1 = \bar{e}_1 + 2\bar{e}_2 \\ \bar{e}'_2 = -3e_1 - 7\bar{e}_2 + \bar{e}_3, \\ \bar{e}'_3 = -2\bar{e}_2 + \bar{e}_3 \end{cases} \text{ será } \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & -3 & 0 \\ 2 & -7 & -2 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

De este último sistema despejamos x_1, x_2 y x_3, para lo que bastará con hallar su matriz inversa; haciendo esto, se obtiene:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 & 3 & 6 \\ -2 & 1 & 2 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{o sea} \quad \begin{cases} x_1 = -5x + 3y + 6z \\ x_2 = -2x + y + 2z \\ x_3 = 2x - y - z \end{cases}$$

5.6. RANGO DE UN SISTEMA DE VECTORES

Cuando hablamos del *rango* por primera vez (véase [015]) lo hicimos sólo para el caso de vectores de $\mathbb{R}^n$ y no se pudo dar una definición plenamente satisfactoria de él, ya que no conocíamos, entonces, nada ni de las bases ni de la dimensión. Vamos ahora a completar lo dicho entonces.

[070]

Se llama *rango* de un sistema $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$, de un número finito de vectores de un cierto espacio, a la dimensión del subespacio que engendra S; el rango se denota poniendo rang S, de modo que:

$$\text{rang } S = \text{dim } [V(S)]$$

El rango de S es, pues, el mayor número de vectores linealmente independientes que hay en S.

OBSERVACIONES

1. Un sistema de vectores $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es linealmente independiente sí y sólo si su rango es igual al número p de vectores que lo forman.
2. En un espacio vectorial de dimensión finita n, un sistema de vectores es generador si y sólo si su rango es n.

CALCULO DEL RANGO

Anteriormente, cuando se estudió el rango en el caso de los vectores de n componentes (véase [015] y [016]), se obtuvo un método práctico para hallarlo, recurriendo a las operaciones elementales. Vamos a ver aquí que calcular el rango de un sistema de p vectores de un espacio vectorial de dimensión n, se reduce a calcular el rango de sus p n-uplas de coordenadas, que son p vectores de n componentes, ya que uno y otro rango son iguales.

[071]

Sea V un espacio vectorial, sobre un cuerpo K, que tiene dimensión n; en V se adopta una base B. Dado un sistema finito $S = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ de vectores de V y si $S^* = (\bar{u}_1^*, \bar{u}_2^*, ..., \bar{u}_p^*)$ es el correspondiente sistema de sus n-uplas de coordenadas (vectores de K^n), se verifica que:

1. $\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p = \bar{o} \Leftrightarrow \lambda_1\bar{u}_1^* + \lambda_2\bar{u}_2^* + \cdots + \lambda_p\bar{u}_p^* = \bar{o}$ $(\lambda_i \in K)$
2. rang S = rang S^*.

DEMOSTRACION

1. Nótese que, para cualquiera que sea el vector $\bar{x} \in V$, se ha denotado por $\bar{x}^*$ al vector $\bar{x}^* = \text{COOR } \bar{x}$, donde COOR es el isomorfismo de coordenadas, de V en K^n (véase [068]), y que, por tanto (por ser COOR lineal):

$$(\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p)^* = \lambda_1\bar{u}_1^* + \lambda_2\bar{u}_2^* + \cdots + \lambda_p\bar{u}_p^*$$

   Nótese, también, que las coordenadas de $\bar{x} \in V$ (vector cualquiera) en la base B son las mismas que las coordenadas de $\bar{x}^* \in K^n$ en la base canónica. En particular, los vectores $\Sigma\lambda_i\bar{u}_i$ y $\Sigma\lambda_i\bar{u}_i^*$ (para $i = 1, 2, ..., p$) tienen las mismas coordenadas, respecto de las referidas bases. La coordenada h-ésima (para $h = 1, 2, ..., n$) de estos dos vectores vale $\Sigma\lambda_i u_{ih}$ (se suma para $i = 1, 2, ..., p$; u_{ih} es la coordenada h-ésima de $\bar{u}_i$). Resulta entonces que tanto la condición $\Sigma\lambda_i\bar{u}_i = \bar{o}$ como la $\Sigma\lambda_i\bar{u}_i^* = \bar{o}$ son, ambas, equivalentes a las n relaciones $\Sigma\lambda_i u_{ih} = 0$ y, por ello, son equivalentes entre sí, como había que comprobar.
2. Según se acaba de probar, entre los vectores de S hay las mismas relaciones de dependencia lineal que entre los vectores de S^*. De ahí que unos ciertos vectores de S son linealmente independientes sí y sólo si lo son los correspondientes vectores de S^* y, por ello, S y S^* tienen el mismo rango.

EJEMPLO

En el espacio vectorial $\mathcal{M}_{2\times3}$, de las matrices reales de tamaño 2×3, consideremos el sistema $S = (M_1, M_2, M_3, M_4)$, formado por las cuatro matrices:

$$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & -1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 2 & -1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \quad M_3 = \begin{bmatrix} 3 & 1 & 2 \\ 4 & 2 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 3 & 2 \\ 0 & 3 & -3 \end{bmatrix}$$

Echando mano de la base usual de $\mathcal{M}_{2\times3}$, los vectores de coordenadas de las matrices dadas son, respectivamente

$$\bar{v}_1 = (1, 0, 0, 2, 1, -1) \qquad \bar{v}_2 = (2, -1, 0, 3, 0, 1)$$
$$\bar{v}_3 = (3, 1, 2, 4, 2, 0) \qquad \bar{v}_4 = (0, 3, 2, 0, 3, -3)$$

El rango del sistema dado S es igual al rango de $S^* = (\bar{v}_1, \bar{v}_2, \bar{v}_3, \bar{v}_4)$. El

rango de este sistema, de vectores de $\mathbb{R}^6$, lo calcularemos recurriendo a las operaciones elementales:

$$\text{rang } S = \text{rang } S^* = \text{rang} \begin{bmatrix} 1 & 0 & 0 & 2 & 1 & -1 \\ 2 & -1 & 0 & 3 & 0 & 1 \\ 3 & 1 & 2 & 4 & 2 & 0 \\ 0 & 3 & 2 & 0 & 3 & -3 \end{bmatrix} = \begin{pmatrix} 2.^a \to 2.^a - 2 \cdot 1.^a \\ 3.^a \to 3.^a - 3 \cdot 1.^a \end{pmatrix}$$

$$= \text{rang} \begin{bmatrix} 1 & 0 & 0 & 2 & 1 & -1 \\ 0 & -1 & 0 & -1 & -2 & 3 \\ 0 & 1 & 2 & -2 & -1 & 3 \\ 0 & 3 & 2 & 0 & 3 & -3 \end{bmatrix} = \begin{pmatrix} 3.^a \to 3.^a + 2.^a \\ 4.^a \to 4.^a + 3 \cdot 2.^a \end{pmatrix}$$

$$= \text{rang} \begin{bmatrix} 1 & 0 & 0 & 2 & 1 & -1 \\ 0 & -1 & 0 & -1 & -2 & 3 \\ 0 & 0 & 2 & -3 & -3 & 6 \\ 0 & 0 & 2 & -3 & -3 & 6 \end{bmatrix} = 3$$

γ SUMA DE SUBESPACIOS

5.7. SUMA Y SUMA DIRECTA

Según ya vimos anteriormente (véase [055]), dados dos subespacios U_1 y U_2, de un cierto espacio vectorial V, el subconjunto $U_1 \cap U_2$ de V es un subespacio de V y, además, $U_1 \cap U_2$ es el subespacio mayor de cuantos están incluidos en U_1 y en U_2. Sin embargo, el subconjunto $U_1 \cup U_2$ no es, en general, un subespacio de V (véase [055]). Así, pues, aunque $U_1 \cup U_2$ es el menor de todos los subconjuntos de V que incluyen a U_1 y a U_2, como no es subespacio, no será el menor de entre los subespacios que incluyen a U_1 y a U_2. Pues bien, en lo que sigue nos vamos a ocupar de dicho subespacio mínimo, que existe, al que se llamará suma de los subespacios U_1 y U_2 y no le denota por $U_1 + U_2$.

SUMA DE DOS SUBESPACIOS

[072]

Sean U_1 y U_2 dos subespacios, de un espacio vectorial V. Se llama suma de U_1 y U_2 al conjunto, que se denota poniendo $U_1 + U_2$, definido mediante

$$U_1 + U_2 = \{\vec{u}_1 + \vec{u}_2 / \vec{u}_1 \in U_1, \vec{u}_2 \in U_2\}$$

Este conjunto es un subespacio de V; es más, se trata del menor de todos los subespacios de V que contienen a U_1 y a U_2.

DEMOSTRACION

El subconjunto $U_1 + U_2$ es subespacio ya que, si $\bar{u}$ y $\bar{v}$ son de $U_1 + U_2$ y para cualesquiera escalares λ y μ, como $\bar{u} = u_1 + \bar{u}_2$ y $\bar{v} = \bar{v}_1 + \bar{v}_2$ para ciertos $\bar{u}_1$, $\bar{v}_1 \in U_1$ y $\bar{u}_2$, $\bar{v}_2 \in U_2$, resulta que

$$\lambda\bar{u} + \mu\bar{v} = (\lambda\bar{u}_1 + \mu\bar{v}_1) + (\lambda\bar{u}_2 + \mu\bar{v}_2) \in U_1 + U_2$$

pues $\lambda\bar{u}_1 + \mu\bar{v}_1 \in U_1$ y $\lambda\bar{u}_2 + \mu\bar{v}_2 \in U_2$.

Todo vector $\bar{u}_1 \in U_1$ es de la suma $U_1 + U_2$, ya que se puede poner $\bar{u}_1 = \bar{u}_1 + \bar{o}$ con $\bar{o} \in U_2$; por tanto, $U_1 \subset U_1 + U_2$. Análogamente, también $U_2 \subset U_1 + U_2$. Para acabar, hemos de comprobar que si U es un subespacio tal que $U_1 \subset U$ y $U_2 \subset U$, entonces $U_1 + U_2 \subset U$; esto es así, ya que:

$$\bar{u} \in U_1 + U_2 \Rightarrow \begin{cases} \bar{u} = \bar{u}_1 + \bar{u}_2 \\ \bar{u}_1 \in U_1,\ \bar{u}_2 \in U_2 \end{cases} \Rightarrow \begin{cases} \bar{u} = \bar{u}_1 + \bar{u}_2 \\ \bar{u}_1,\ \bar{u}_2 \in U \end{cases} \Rightarrow \bar{u} \in U$$

SUMA DE VARIOS SUBESPACIOS

[073]

Sean $U_1, U_2, \ldots, U_p$ subespacios, de un espacio vectorial V. Se llama suma de estos subespacios al siguiente conjunto de vectores:

$$\sum_{i=1}^{p} U_i \equiv U_1 + U_2 + \cdots + U_p = \{\bar{u}_1 + \bar{u}_2 + \cdots + \bar{u}_p / \bar{u}_i \in U_i \text{ para } i = 1, 2, \ldots, p\}$$

Este conjunto es un subespacio de V; es más, se trata del menor de todos los subespacios de V que contienen a todos los U_i ($i = 1, 2, \ldots, p$).

DEMOSTRACION

Siguiendo los mismos pasos de la demostración anterior, se comprueba trivialmente la veracidad de cuanto se acaba de decir. Así, por ejemplo, para ver que ΣU_i ($i = 1, 2, \ldots, p$) es un subespacio, se tiene:

$$\begin{pmatrix} \bar{u},\ \bar{v} \in \Sigma U_i \\ \lambda,\ \mu \text{ escalares} \end{pmatrix} \Rightarrow \begin{pmatrix} \bar{u} = \Sigma \bar{u}_i,\ \bar{v} = \Sigma \bar{v}_i \\ \bar{u}_i,\ \bar{v}_i \in U_i\ (i = 1, 2, \ldots, p) \\ \lambda,\ \mu \text{ escalares} \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{pmatrix} \lambda\bar{u} + \mu\bar{v} = \Sigma(\lambda\bar{u}_i + \mu\bar{v}_i) \\ \lambda\bar{u}_i + \mu\bar{v}_i \in U_i\ (i = 1, 2, \ldots, p) \end{pmatrix} \Rightarrow \lambda\bar{u} + \mu\bar{v} \in \Sigma U_i$$

EJEMPLOS

1. En el espacio vectorial $V = \mathbb{R}^4$, consideremos los subespacios

$$U_1 = \{(\alpha, \beta, 0, 0)/\alpha, \beta \in \mathbb{R}\} \quad \text{y} \quad U_2 = \{(0, \lambda, \mu, 0)/\lambda, \mu \in \mathbb{R}\}$$

La suma de estos dos subespacios está formada por los vectores $(\alpha, \beta + \lambda, \mu, 0)$ para α, β, λ y μ recorriendo $\mathbb{R}$. Resulta entonces evidente que dicha suma es:

$$U_1 + U_2 = \{(x, y, z, 0)/x, y, z \in \mathbb{R}\}$$

Obsérvese que cualquier vector de la suma se puede obtener, como suma de un vector de U_1 y otro de U_2, de muchas maneras; así, por ejemplo:

$$(2, 3, -1, 0) = (2, \beta, 0. 0) + (0, \lambda, -1, 0)$$

para cualesquiera β, $\lambda \in \mathbb{R}$ tales que $\beta + \lambda = 3$.

2. En el mismo espacio $V = \mathbb{R}^4$, consideremos ahora los subespacios:

$$U_1' = \{(\alpha, 0, 0, 0)/\alpha \in \mathbb{R}\} \quad \text{y} \quad U_2 = \{(0, \lambda, \mu, 0)/\lambda, \mu \in \mathbb{R}\}$$

La suma $U_1' + U_2$ es, evidentemente:

$$U_1' + U_2 = \{(x, y, z, 0)/x, y, z \in \mathbb{R}\}$$

que es la misma que en el caso anterior. Entre estos dos ejemplos hay una notable diferencia: contrariamente a lo que ocurría antes, cualquier vector de la suma $U_1' + U_2$ se puede expresar de una única manera como suma de un vector de U_1' y otro de U_2; esta descomposición es la:

$$(x, y, z, 0) = (x, 0, 0, 0) + (0, y, z, 0)$$

Pues bien, cuando, como aquí, la descomposición sea única, la suma de subespacios se llama suma directa.

SUMA DIRECTA

[074]

Sean $U_1, U_2, ..., U_p$ subespacios de un espacio vectorial V.

Definición.—Se dice que $U_1, U_2, ..., U_p$ son *subespacios independientes* o que $U_1 + U_2 + \cdots + U_p$ es una *suma directa* de subespacios y se la escribe en la forma $U_1 \oplus U_2 \oplus \cdots \oplus U_p$, si cualquier vector de dicha suma de subespacios puede expresarse de una única forma como suma de vectores de $U_1, U_2, ..., U_p$; esto es, si:

$$\left.\begin{array}{l} \vec{u}_1 + \vec{u}_2 + \cdots + \vec{u}_p = \vec{u}_1' + \vec{u}_2' + \cdots + \vec{u}_p' \\ \vec{u}_i, \vec{u}_i' \in U_i \quad \text{para} \quad i = 1, 2, ..., p \end{array}\right\} \Rightarrow \vec{u}_i = \vec{u}_i' \ (i = 1, 2, ..., p) \ \text{(I)}$$

Caracterización.—La suma $U_1 + U_2 + \cdots + U_p$ es directa si y sólo si

$$\left.\begin{array}{l} \vec{u}_1 + \vec{u}_2 + \cdots + \vec{u}_p = \vec{0} \\ \vec{u}_i \in U_i \ (i = 1, 2, ..., p) \end{array}\right\} \Rightarrow \vec{u}_i = \vec{0} \quad (\text{para } i = 1, 2, ..., p) \quad \text{(II)}$$

DEMOSTRACION

1. Veamos que (I) implica (II). La relación $\bar{u}_1 + \bar{u}_2 + \cdots + \bar{u}_p = \bar{o}$ se puede expresar en la forma:

$$\bar{u}_1 + \bar{u}_2 + \cdots + \bar{u}_p = \bar{o} + \bar{o} + \cdots + \bar{o} \quad \text{donde} \quad \bar{u}_i \in U_i \quad \text{y} \quad \bar{o} \in U_i$$

y entonces, aplicando (I) a este caso en el que $\bar{u}_i' = \bar{o}$, obtenemos que $\bar{u}_i = \bar{o}$ para $i = 1, 2, ..., p$, como había que probar.

2. Veamos ahora que (II) implica (I). Si se verifica que $\Sigma \bar{u}_i = \Sigma \bar{u}_i'$ con $\bar{u}_i$, $\bar{u}_i' \in U_i$ (para $i = 1, 2, ..., p$), se verificará también que $\Sigma (\bar{u}_i - \bar{u}_i') = \bar{o}$ con $\bar{u}_i - \bar{u}_i' \in U_i$ (para $i = 1, 2, ..., p$) y entonces, aplicando (I) a este caso, se concluye que $\bar{u}_i - \bar{u}_i' = \bar{o}$, o sea $\bar{u}_i = \bar{u}_i'$ para $i = 1, 2, ..., p$, como había que comprobar.

EJERCICIO

Comprobar que si la suma de varios subespacios es directa, entonces también es directa la suma de algunos (cualesquiera) de ellos. Esto es:

$$\begin{pmatrix} U_1 + \cdots + U_h + \cdots + U_p \\ \text{es suma directa} \end{pmatrix} \Rightarrow \begin{pmatrix} U_1 + \cdots + U_h \\ \text{es directa} \end{pmatrix}$$

RESOLUCION

Hay que comprobar que, si $\bar{u}_1 + \cdots + \bar{u}_h = \bar{o}$ para ciertas $\bar{u}_1 \in U_1$, ..., $\bar{u}_h \in U_n$, entonces es $\bar{u}_1 = \cdots = \bar{u}_h = \bar{o}$. Así ocurre, ya que como la suma de todos, los p, subespacios es directa, se verifica que:

$$\begin{pmatrix} \bar{u}_1 + \cdots + \bar{u}_h = \bar{o} \\ \bar{u}_1 \in U_1, \ ..., \ \bar{u}_h \in U_h \end{pmatrix} \Rightarrow \begin{pmatrix} \bar{u}_1 + \cdots + \bar{u}_h + \bar{o} + \cdots + \bar{o} = \bar{o} \\ \bar{u}_1 \in U_1, \ ..., \ \bar{u}_h \in U_h, \ \bar{o} \in U_{h+1}, \ ..., \ \bar{o} \in U_p \end{pmatrix} \Rightarrow$$

$$\Rightarrow \quad \bar{u}_1 = \bar{o}, \ ..., \ \bar{u}_h = \bar{o}, \ \bar{o} = \bar{o}, \ ..., \ \bar{o} = \bar{o}$$

EJEMPLO

Sea V el espacio vectorial de los polinomios reales (con una indeterminada x). Sean U_1 y U_2 los subespacios:

$$U_1 = \{p(x) \in V / p(1) = p(-1) = 0\} = \{c(x)(x^2 - 1) / c(x) \in V\}$$

$$U_2 = \{\text{polinomios de grado menor que } 2\} = \{a + bx / a, b \in \mathbb{R}\}$$

(nótese que los polinomios $p(x)$ tales que $p(1) = 0$ y $p(-1) = 0$ son los divisibles por $x - 1$ y por $x + 1$, es decir, los de la forma $c(x)(x - 1)(x + 1)$, donde

$c(x)$ es un polinomio). La suma $U_1 + U_2$ es todo el espacio V ya que cualquier polinomio $p(x) \in V$ se puede poner en la forma

$$p(x) = c(x)(x^2 - 1) + (a + bx)$$

donde $c(x)$ y $a + bx$ son el cociente y el resto de la división de $p(x)$ entre $x^2 - 1$. Como esta división conduce a unos únicos cociente y resto, la anterior descomposición es única y, por ello, la suma $U_1 + U_2$ es directa.

SUMA DIRECTA DE DOS SUBESPACIOS

Además de particularizar, para el caso $p = 2$, lo ya dicho para la suma de p subespacios, estudiaremos aquí una caracterización típica de la suma directa de dos subespacios, que no es válida para $p > 2$.

[075]

Sean U_1 y U_2 dos subespacios de un espacio vectorial V. De acuerdo con la definición general, $U_1 + U_2$ será *suma directa*, lo que se denota poniendo $U_1 \oplus U_2$, si se verifica una cualquiera de las dos condiciones siguientes (I) o (II), equivalentes entre sí:

$$\bar{u}_1 + \bar{u}_2 = \bar{u}_1' + \bar{u}_2' \Rightarrow [\bar{u}_1 = \bar{u}_1' \text{ y } \bar{u}_2 = \bar{u}_2'] \quad \text{(I)}$$

$$\bar{u}_1 + \bar{u}_2 = \bar{0} \Rightarrow \bar{u}_1 = \bar{u}_2 = \bar{0} \quad \text{(II)}$$

(donde $\bar{u}_1, \bar{u}_1' \in U_1$ y $\bar{u}_2, \bar{u}_2' \in U_2$).

Los subespacios U_1 y U_2 son independientes si y sólo si su intersección es nula; esto es:

$$[U_1 + U_2 \text{ directa}] \Leftrightarrow [U_1 \cap U_2 = O] \quad \text{(III)}$$

DEMOSTRACION

La equivalencia de (I) y (II) ya se probó en [074]; probemos la (III):

1. Comprobemos la implicación $\Rightarrow$ viendo que (si $U_1 + U_2$ es directa), de ser $\bar{u} \in U_1 \cap U_2$, se verifica que $\bar{u} = \bar{0}$. Como $\bar{u} + (-\bar{u}) = \bar{0}$ y $\bar{u} \in U_1$ y $-\bar{u} \in U_2$, de la condición (II) se deduce que $\bar{u} = \bar{0}$. como había que comprobar.
2. Comprobemos la implicación $\Leftarrow$ viendo que (si $U_1 \cap U_2 = O$), de ser $\bar{u}_1 + \bar{u}_2 = \bar{0}$ con $\bar{u}_1 \in U_1$ y $\bar{u}_2 \in U_2$, se verifica que $\bar{u}_1 = \bar{u}_2 = \bar{0}$. Así ocurre, ya que como $\bar{u}_1 = -\bar{u}_2$ donde $\bar{u}_1 \in U_1$ y $-\bar{u}_2 \in U_2$, resulta que $\bar{u}_1$ y $\bar{u}_2$ son, ambos de U_1 y de U_2, luego $\bar{u}_1$ y $\bar{u}_2$ son nulos por pertenecer a $U_1 \cap U_2 = O$.

OBSERVACION

Si la suma de varios subespacios $U_1, U_2, \ldots, U_p$ es directa, entonces la intersección de cada dos de ellos es nula, $U_i \cap U_j = O$ para $i \neq j$, pero el recíproco es falso (en general). En efecto:

- Si la suma $U_1 + U_2 + \cdots + U_p$ es directa, entonces también lo son todas las sumas $U_i + U_j$ para $i \neq j$ (con $i, j = 1, 2, ..., p$) y, según se acaba de comprobar, de ello se deduce que $U_i \cap U_j = O$.
- Si $U_i \cap U_j = O$ para cualesquiera $i \neq j$, puede ocurrir sin embargo que la suma $U_1 + U_2 + \cdots + U_p$ no sea directa. Así ocurre, por ejemplo, en el espacio vectorial $\mathbb{R}^3$, con los siguientes subespacios U_1, U_2 y U_3:

$$U_1 = \{(\alpha, \beta, 0)/\alpha, \beta \in \mathbb{R}\} \quad ; \quad U_2 = \{(0, 0, \gamma)/\gamma \in \mathbb{R}\}$$
$$U_3 = \{(\lambda, \lambda, \lambda)\ \lambda \in \mathbb{R}\}$$

Es evidente que $U_1 \cap U_2 = U_2 \cap U_3 = U_3 \cap U_1 = O$ y que $U_1 + U_2 + U_3$ no es directa por no ser única la descomposición de un vector de dicha suma:

$$(\alpha, \beta, 0) + (0, 0, \gamma) + (\lambda, \lambda, \lambda) = (\alpha + h, \beta + h, 0) + (0, 0, \gamma + h) + \\ + (\lambda - h, \lambda - h, \lambda - h)$$

SUMA DE SUBESPACIOS UNIDIMENSIONALES

[076]

Sea V un espacio vectorial y sea $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ un sistema finito de vectores de V. Se verifica que:

1. **El subespacio que engendra $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es igual a la suma de los p subespacios que engendran los vectores del sistema; esto es:**

$$\mathcal{V}(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p) = \mathcal{V}(\bar{u}_1) + \mathcal{V}(\bar{u}_2) + \cdots + \mathcal{V}(\bar{u}_p)$$

2. **El sistema $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ de vectores no nulos es linealmente independiente sí y sólo si los subespacios que engendran sus p vectores tienen suma directa; esto es:**

$$\begin{pmatrix} (\bar{u}_1, \bar{u}_2, ..., \bar{u}_p) \\ \text{independiente} \end{pmatrix} \Leftrightarrow \begin{pmatrix} \mathcal{V}(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p) = \\ = \mathcal{V}(\bar{u}_1) \oplus \mathcal{V}(\bar{u}_2) \oplus \cdots \oplus \mathcal{V}(\bar{u}_p) \end{pmatrix}$$

DEMOSTRACION

1. La igualdad a demostrar es evidente pues los subespacios de su primer y de su segundo miembros están formados, ambos, por los mismos vectores, que son los:

$$\lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_p \bar{u}_p \quad (\lambda_1, \lambda_2, ... \lambda_p \text{ escalares cualesquiera})$$

2. La suma de los $\mathcal{V}(\bar{u}_i)$ es directa si y sólo si:

$$\left.\begin{array}{l} \lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_p \bar{u}_p = \bar{o} \\ \lambda_1, \lambda_2, ..., \lambda_p \text{ escalares} \end{array}\right\} \Rightarrow \lambda_1 \bar{u}_1 = \lambda_2 \bar{u}_2 = \cdots = \lambda_p \bar{u}_p = \bar{o}$$

Ahora bien como todos los vectores $\bar{u}_1, \bar{u}_2, ..., \bar{u}_p$ son no nulos, resulta que:

$$\lambda_1\bar{u}_1 = \lambda_2\bar{u}_2 = \cdots = \lambda_p\bar{u}_p = \bar{o} \quad \Leftrightarrow \quad \lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$$

luego la suma de los subespacios $\mathcal{V}(\bar{u}_i)$ es directa si y sólo si el sistema $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es independiente, como había que comprobar.

SUBESPACIOS SUPLEMENTARIOS

[077]

En un espacio vectorial V, dos subespacios U_1 y U_2 se dicen *suplementarios* si cualquier vector $\bar{u} \in V$ se puede expresar de una sola manera como suma, $\bar{u} = \bar{u}_1 + \bar{u}_2$, de un vector $\bar{u}_1 \in U_1$ y otro $\bar{u}_2 \in U_2$. Según lo que se acaba de decir acerca de la suma directa de dos subespacios, se verifica que:

$$\left(\begin{array}{c}\text{Los subespacios}\\ U_1 \text{ y } U_2 \text{ de } V \text{ son}\\ \text{suplementarios}\end{array}\right) \Leftrightarrow U_1 \oplus U_2 = V \Leftrightarrow \left(\begin{array}{l}U_1 + U_2 = V\\ U_1 \cap U_2 = O\end{array}\right)$$

EJEMPLO

Sea V el espacio vectorial de las funciones de $\mathbb{R}$ en $\mathbb{R}$. Se dice que una función $f \in V$ es par si se verifica que $f(-x) = f(x)$ para todo $x \in \mathbb{R}$; se dice que $f \in V$ es impar si $f(-x) = -f(x)$ para todo $x \in \mathbb{R}$. Es fácil comprobar que

$$U_1 = \{f \in V / f \text{ es par}\} \quad \text{y} \quad U_2 = \{f \in V / f \text{ es impar}\}$$

son dos subespacios de V. La única función que es par e impar es la función nula, ya que debe ser $f(x) = -f(x)$ para todo $x \in \mathbb{R}$; por tanto $U_1 \cap U_2 = O$. Por otra parte, toda función $f \in V$ se puede poner en la forma:

$$f(x) = f_1(x) + f_2(x)$$

siendo

$$f_1(x) = \frac{f(x) + f(-x)}{2} \quad \text{y} \quad f_2(x) = \frac{f(x) - f(-x)}{2}$$

y estas dos funciones f_1 y f_2 son par e impar, respectivamente; por ello, resulta que $U_1 + U_2 = V$. Resumiendo: U_1 y U_2 son subespacios suplementarios de V.

En un espacio vectorial V, se dice que un subespacio U es *máximal* si es el suplementario de algún subespacio unidimensional. Un subespacio U es máximal si está estrictamente incluido en V (o sea, $U \subset V$ y $U \neq V$) y no hay ningún subespacio que le incluya estrictamente, salvo el propio V.

5.8. DIMENSION DE LA SUMA

DIMENSIONES Y BASES DE LOS SUBESPACIOS SUPLEMENTARIOS

[078]

En un espacio vectorial V de dimensión finita, se verifica que:

1. $\left(\begin{array}{c} U_1 \text{ y } U_2 \text{ subespacios} \\ \text{suplementarios de } V. \\ B_1 \text{ es una base de } U_1; \\ B_2 \text{ es una base de } U_2 \end{array}\right) \Rightarrow \left(\begin{array}{c} B = B_1 \cup B_2 \text{ es} \\ \text{una base de } V \end{array}\right)$

2. $\left(\begin{array}{c} U_1 \text{ y } U_2 \text{ subespacios} \\ \text{suplementarios de } V \end{array}\right) \Rightarrow \dim U_1 + \dim U_2 = \dim V$

3. $\left(\begin{array}{l} B = (\bar{e}_1, \ldots, \bar{e}_s, \bar{e}_{s+1}, \ldots, \bar{e}_n) \text{ base de } V \\ B_1 = (\bar{e}_1, \ldots, \bar{e}_s), \quad U_1 = \mathcal{V}(B_1) \\ B_2 = (\bar{e}_{s+1}, \ldots, \bar{e}_n), \quad U_2 = \mathcal{V}(B_2) \end{array}\right) \Rightarrow \left(\begin{array}{c} U_1 \text{ y } U_2 \\ \text{subespacios} \\ \text{suplementarios} \\ \text{de } V \end{array}\right)$

4. Todos subespacio de V tiene algún subespacio suplementario.

DEMOSTRACION

1. Hay que comprobar que cualquier $\bar{u} \in V$ puede expresarse de una sola manera como combinación lineal de los vectores de B_1 y B_2. Como $V = U_1 \oplus U_2$, se sabe que $\bar{u}$ se puede expresar de una sola manera como suma $\bar{u} = \bar{u}_1 + \bar{u}_2$ con $\bar{u}_1 \in U_1$ y $u_2 \in U_2$. Como B_1 y B_2 son bases de U_1 y U_2, los vectores $\bar{u}_1$ y $\bar{u}_2$ se pueden poner de manera única como combinaciones lineales de los vectores de B_1 y B_2, respectivamente. Consecuencia de todo ello es que $\bar{u}$ se puede poner como combinación lineal de los vectores de B y que esto se puede hacer de una sola manera.
2. Esta propiedad es consecuencia de la anterior, ya que las dimensiones de V, U_1 y U_2 son los números n, n_1 y n_2 de vectores que tienen B, B_1 y B_2, respectivamente, luego $n = n_1 + n_2$ ya que $B = B_1 \cup B_2$ y $B_1 \cap B_2 = \varnothing$.
3. Hay que comprobar que $U_1 + U_2 = V$ y que $U_1 \cap U_2 = O$.

   - Veamos que $U_1 + U_2 = V$. Como B es base de V, cualquiera $\bar{u} \in V$ se puede expresar (para ciertos escalares λ_i) en la forma

   $$\bar{u} = (\lambda_1 \bar{e}_1 + \cdots + \lambda_s \bar{e}_s) + (\lambda_{s+1} \bar{e}_{s+1} + \cdots + \lambda_n \bar{e}_n)$$

   donde el vector del primer paréntesis es de U_1 y el del segundo es de U_2, luego $\bar{u} \in U_1 + U_2$. En consecuencia $U_1 + U_2 = V$.

- Veamos que $U_1 \cap U_2 = O$. Si $\bar{u} \in U_1 \cap U_2$, recurriendo a que B_1 y B_2 son bases de U_1 y de U_2, $\bar{u}$ se podrá expresar recurriendo a una y otra base, con lo que para ciertos escalares λ_i será:

$$\left.\begin{aligned} \lambda_1\bar{e}_1 + \cdots + \lambda_s\bar{e}_s = \bar{u} \\ \lambda_{s+1}\bar{e}_{s+1} + \cdots + \lambda_n\bar{e}_n = \bar{u} \end{aligned}\right\}$$

luego

$$\lambda_1\bar{e}_1 + \cdots + \lambda_s\bar{e}_s - \lambda_{s+1}\bar{e}_{s+1} - \cdots - \lambda_n\bar{e}_n = \bar{o}$$

Ahora bien, como B es base de V, de la última relación se deduce que todos los escalares λ_i son nulos y en consecuencia es $\bar{u} = \bar{o}$, o sea, el único vector de $U_1 \cap U_2$ es el $\bar{o}$, como había que comprobar.

4. Si el subespacio es todo V o si es O, entonces O o V es, respectivamente, su suplementario. Si el subespacio es propio, entonces tiene alguna base; sea B_1 una tal base. Completemos B_1 hasta obtener una base de V y llamemos B_2 al sistema de vectores que se han añadido. Según la propiedad anterior, el subespacio $V(B_2)$ es suplementario del dado.

EJEMPLO

En $\mathbb{R}^3$, el siguiente subconjunto U_1 es suplementario de U_2 y también lo es del U_2':

$$U_1 = \{(x, y, 0)/x, y \in \mathbb{R}\} \quad ; \quad U_2 = \{(0, 0, z)/z \in \mathbb{R}\}$$
$$U_2' = \{(\alpha, \alpha, \alpha)/\alpha \in \mathbb{R}\}$$

Una base B_1 de U_1 está formada por (1, 0, 0) y (0, 1, 0); una base B_2 de U_2 está formada por (0, 0, 1); una base B_2' de U_2' está formada por (1, 1, 1). Nótese que $B_1 \cup B_2$ y $B_1 \cup B_2'$ son, las dos, bases de $\mathbb{R}^3$.

FORMULA DE LAS DIMENSIONES (O DE GRASSMANN)

[079]

Si U_1 y U_2 son dos subespacios de un espacio vectorial V de dimensión finita, se verifica que

$$\dim U_1 + \dim U_2 = \dim (U_1 + U_2) + \dim U_1 \cap U_2$$

DEMOSTRACION

Sea $B_0 = (e_1, e_2, \ldots, e_m)$ una base de $U_1 \cap U_2$. Sabemos (véase [065]) que esta base se puede completar hasta obtener una base de U_1 y que también es posible

ampliarla hasta conseguir una base de U_2; sean B_1 y B_2 dichas bases:

$$B_1 = B_0 \cup S_1 = (e_1, ..., e_m) \cup (\bar{u}_1, ..., \bar{u}_r) \quad \text{base de } U_1$$
$$B_2 = B_0 \cup S_2 = (\bar{e}_1, ..., \bar{e}_m) \cup (\bar{v}_1, ..., \bar{v}_s) \quad \text{base de } U_2$$

Vamos a comprobar que $B = B_0 \cup S_1 \cup S_2$ es una base de $U_1 + U_2$ con lo que, como consecuencia, se habrá probado la fórmula de las dimensiones ya que sería

$$\dim U_1 = m + r \qquad \dim (U_1 + U_2) = m + r + s$$
$$\dim U_2 = m + s \qquad \dim (U_1 \cap U_2) = m$$

- Veamos primeramente que B es sistema generador de $U_1 + U_2$. Cualquier vector $\bar{u}$ de esta suma se puede expresar en la forma $\bar{u} = \bar{u}_1 + \bar{u}_2$ con $\bar{u}_1 \in U_1$ y $\bar{u}_2 \in U_2$; como B_1 y B_2 son bases de U_1 y U_2, resulta que $\bar{u}_1$ y $\bar{u}_2$ son combinaciones lineales de los vectores de $B_0 \cup S_1$ y de $B_0 \cup S_2$, respectivamente, con lo que $\bar{u} = \bar{u}_1 + \bar{u}_2$ es combinación lineal de los vectores de $B_0 \cup S_1 \cup S_2 = B$, como había que comprobar.
- Veamos ahora que B es un sistema independiente. Supongamos, para ello, que existen ciertas escalares λ_k, α_i y β_j tales que:

$$\Sigma \lambda_k \bar{e}_k + \Sigma \alpha_i \bar{u}_i + \Sigma \beta_j \bar{v}_j = \bar{o} \quad (k = 1, ... m;\ i = 1, ..., r;\ j = 1, ..., s)$$

Llamando $\bar{w} = \Sigma \lambda_k \bar{e}_k + \Sigma \alpha_i \bar{u}_i$, que depende linealmente de $B_1 = B_0 \cup S_1$ y por ello pertenece a U_1, nos encontramos con que $\bar{w} = -\Sigma \beta_j \bar{v}_j$ y entonces $\bar{w} \in U_2$, pues depende linealmente de $S_2 \subset B_2$. Por tanto, resulta que $\bar{w} \in U_1 \cap U_2$, luego $\bar{w}$ es una combinación lineal de B_0, o sea $\bar{w} = \Sigma \mu_h \bar{e}_h$ (para $h = 1, ..., m$) para ciertos escalares μ_h. Por tanto, $\Sigma \mu_h \bar{e}_h + \Sigma \beta_j \bar{v}_j = \bar{o}$ y, como $B_2 = B_0 \cup S_2$ es sistema independiente, de aquí se deduce que $\mu_h = 0$ para $h = 1, ..., m$ y que $\beta_j = 0$ para $j = 1, ..., s$. Como todos los β_j son nulos, resulta que $\Sigma \lambda_k \bar{e}_k + \Sigma \alpha_i \bar{u}_i = \bar{o}$ y, como $B_1 = B_0 \cup S_1$ es sistema independiente, de esta igualdad se deduce que son nulos todos los escalares λ_k y α_i. Resumiendo: si una combinación lineal de vectores de B es nula, ha de tener nulos todos sus coeficientes, como había que comprobar.

CAPITULO

6

Aplicaciones lineales

Después de haber hablado, en el capítulo anterior, sobre los espacios vectoriales, procede que ahora nos ocupemos de las aplicaciones entre ellos que «conservan» la suma y el producto por un escalar. Estas aplicaciones, que se llaman homomorfismos (de homo, semejante o igual, y morfomo, forma) o aplicaciones lineales, son aquellas en las que la suma de dos vectores se transforma en la suma de sus transformados y el doble, el triple, etc., de un vector se transforma en el doble, el triple, etc., de su transformado. Podemos decir, pues, que las aplicaciones lineales son aquellas que «respetan» la estructura de espacio vectorial.

Hay un gran número de fenómenos que, al menos en primera aproximación, se conducen de modo lineal. Así ocurre, por ejemplo, con las deformaciones elásticas de un cuerpo sometido a ciertas fuerzas: si cuando las fuerzas que actúan en un punto son $\vec{F}$ y $\vec{F}'$ se producen, respectivamente, unas deformaciones $\vec{D}$ y $\vec{D}'$, se supone que, para deformaciones pequeñas, cuando actúen las fuerzas $k\vec{F}$ y $\vec{F}+\vec{F}'$ se producirán las deformaciones $k\vec{D}$ y $\vec{D}+\vec{D}'$, respectivamente. Esta hipótesis, que algunos llaman «principio de superposición» y que es una buena aproximación de la realidad, significa que la aplicación $\vec{F} \longmapsto \vec{D}$, que a cada fuerza le asocia la deformación que produce, es lineal entre espacios vectoriales.

α APLICACIONES LINEALES

6.1. DEFINICION Y PROPIEDADES

HOMOMORFISMOS O APLICACIONES LINEALES

[080]

Dados dos espacios vectoriales V y W, ambos con el mismo cuerpo de escalares K, y una aplicación $f: V \to W$, se dice que f es un *homomorfismo* o *aplicación lineal* si para cualesquiera $\bar{u}, \bar{v} \in V$ y $\lambda, \mu \in K$ se verifica:

$$f(\lambda\bar{u} + \mu\bar{v}) = \lambda f(\bar{u}) + \mu f(\bar{v})$$

Es obvio que esta condición equivale a que, para cualesquiera $\bar{u}, \bar{v} \in V$ y todo $\lambda \in K$, sea:

$$f(\bar{u} + \bar{v}) = f(\bar{u}) + f(\bar{v}) \quad \text{y} \quad f(\lambda\bar{u}) = \lambda f(\bar{u})$$

Una aplicación lineal $f: V \to W$ en la que $W = V$ recibe el nombre de *endomorfismo*.

EJEMPLOS

1. La aplicación $f: \mathbb{R}^3 \to \mathbb{R}^2$ definida mediante

$$f(x, y, z) = (2x - y + 4z, -3x + 5y + 6z)$$

es lineal, del espacio vectorial $\mathbb{R}^3$ en el $\mathbb{R}^2$.
En general, también es lineal la aplicación $f: \mathbb{R}^n \to \mathbb{R}^m$ en la que:

$$f(x_1, x_2, \ldots, x_n) = \left(\sum_{i=1}^{n} a_i x_i, \sum_{i=1}^{n} b_i x_i, \overset{m}{\ldots}, \sum_{i=1}^{n} l_i x_i \right)$$

2. Si U_1 y U_2 son dos subespacios suplementarios de un espacio vectorial V, o sea, si $V = U_1 \oplus U_2$, se llaman *proyecciones* a las aplicaciones:

$$p_1: V \to U_1 \quad ; \quad p_2: V \to U_1$$
$$\bar{u} \mapsto \bar{u}_1 \qquad \bar{u} \mapsto \bar{u}_2$$

donde $\bar{u} = \bar{u}_1 + \bar{u}_2$ es la descomposición de un vector $\bar{u} \in V$ es suma de $\bar{u}_1 \in U_1$ y $\bar{u}_2 \in U_2$. Estas proyecciones son, ambas, aplicaciones lineales.

3. Sea $V = \mathcal{C}[0, 1]$ el espacio vectorial de las funciones reales que son continuas en el intervalo [0, 1]. La aplicación

$$\mathcal{J}: V \to \mathbb{R} \quad , \quad \varphi \mapsto \mathcal{J}(\varphi) = \int_0^1 \varphi(x)\, dx$$

es lineal entre espacios vectoriales, ya que

$$\mathcal{J}(\lambda\varphi + \mu\psi) = \int_0^1 (\lambda\varphi(x) + \mu\psi(x))\,dx =$$

$$= \lambda \int_0^1 \varphi(x)\,dx + \mu \int_0^1 \psi(x)\,dx = \lambda\mathcal{J}(\varphi) + \mu\mathcal{J}(\psi)$$

4. De entre todas las posibles aplicaciones lineales de un espacio V en otro W, hay dos que son triviales; se trata de las siguientes:

   - La *aplicación nula* $o: V \to W$, en la que $o(\bar{u}) = \bar{o}$ para todo $\bar{u} \in V$.
   - La *aplicación identidad* $i: V \to W$, en la que $i(\bar{u}) = \bar{u}$ para todo $\bar{u} \in V$.

PRIMERAS PROPIEDADES

[081]

Sea $f: V \to W$ una aplicación lineal entre espacios vectoriales, sobre el mismo cuerpo K. Se verifica que:

1. $f\left(\sum_{i=1}^{p} \lambda_i \bar{u}_i\right) = \sum_{i=1}^{p} \lambda_i f(\bar{u}_i)$ $\quad (\forall\, \bar{u}_i \in V \text{ y } \forall\, \lambda_i \in K \text{ para } i = 1, 2, \ldots, p)$

   En particular $f(\bar{o}) = \bar{o}$ y $f(-\bar{u}) = -f(\bar{u})$ para cualquier $\bar{u} \in V$.

2. Si $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_p)$ es un sistema dependiente de vectores de V, entonces $(f(\bar{u}_1), f(\bar{u}_2), \ldots, f(\bar{u}_p))$ es un sistema dependiente de W.

3. Si, además de que $f: V \to W$ es lineal, también se sabe que $g: W \to U$ es aplicación lineal, entonces su composición $g \circ f: V \to U$ es, a su vez, una aplicación lineal.

NOTA: si $(\bar{u}_1, \ldots, \bar{u}_p)$ es un sistema independiente de V, el sistema $(f(\bar{u}_1), \ldots, f(\bar{u}_p))$ de W puede ser dependiente o independiente, según los casos.

DEMOSTRACION

1. Este resultado se obtiene aplicando reiteradamente la definición de aplicación lineal. Tomando todos los λ_i nulos, se obtiene $f(\bar{o}) = \bar{o}$. Tomando $\lambda_1 = -1$ y los demás escalares nulos, se llega a $f(-\bar{u}) = -f(\bar{u})$.

2. Sabemos que hay unos ciertos escalares $\lambda_1, \lambda_2, \ldots, \lambda_p$ no todos nulos tales que:

$$\bar{o} = \lambda_1 u_1 + \lambda_2 \bar{u}_2 + \cdots + \lambda_p \bar{u}_p$$

Por tanto, como f es lineal, se verificará que:

$$\bar{o} = f(\bar{o}) = f\left(\sum_{i=1}^{p} \lambda_i \bar{u}_i\right) = \lambda_i f(\bar{u}_1) + \lambda_2 f(\bar{u}_2) + \cdots + \lambda_p f(\bar{u}_p)$$

y, por tanto, $f(\bar{u}_1), f(\bar{u}_2), \ldots, f(\bar{u}_p)$ forman un sistema dependiente.

3. Para cualesquiera $\bar{u}_1, \bar{u}_2 \in V$ y $\lambda_1, \lambda_2 \in K$, como f y g son ambas lineales, se verifica que

$$(g \circ f)(\lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2) = g[f(\lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_2)] = g[\lambda_1 f(\bar{u}_1) + \lambda_2 f(\bar{u}_2)] =$$

$$= \lambda_1 g(f(\bar{u}_1)) + \lambda_2 g(f(\bar{u}_2)) = \lambda_1 (g \circ f)(\bar{u}_1) + \lambda_2 (g \circ f)(\bar{u}_2)$$

por lo que $g \circ f$ es efectivamente lineal.

EJEMPLO

Considérese la aplicación lineal $f: \mathbb{R}^3 \to \mathbb{R}^2$ definida por:

$$f(x, y, z) = (x, y)$$

Los vectores $\bar{u} = (1, 0, 0)$ y $\bar{v} = (0, 1, 0)$ son linealmente independientes y también lo son sus imágenes $f(\bar{u}) = (1, 0)$ y $f(\bar{v}) = (0, 1)$. Sin embargo, los vectores $\bar{a} = (1, 0, 1)$ y $\bar{b} = (2, 0, 0)$ también son independientes, pero sus imágenes $f(\bar{a}) = (1, 0)$ y $f(\bar{b}) = (2, 0)$ son dependientes. De este ejemplo se desprende que, en general, la independencia de vectores no es una propiedad que se conserve por las aplicaciones lineales.

NUCLEO E IMAGEN; RANGO

Asociados a toda aplicación lineal $f: V \to W$, hay dos subespacios que desempeñan un papel destacado; se trata del núcleo, subespacio de V que forman los vectores que tienen a $\bar{o}$ por imagen, y del subespacio $f(V)$ de W, al que se llama imagen de f.

[082]

Si $f: V \to W$ es una aplicación lineal entre espacios vectoriales, sobre un mismo cuerpo K, entonces:

1. El conjunto imagen $f(V)$ es un subespacio vectorial de W, que se llama *imagen* de la aplicación lineal f y se denota por $\text{Im}(f)$.
2. Si $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es un sistema generador de V, entonces $(f(\bar{u}_1), f(\bar{u}_2), ..., f(\bar{u}_p))$ es un sistema generador de $f(V) = \text{Im}(f)$. Se llama *rango* de f a:

$$\text{rang } f = \dim(\text{Im}(f)) = \text{rang}(f(\bar{u}_1), f(\bar{u}_1), f(\bar{u}_2), ..., f(\bar{u}_p))$$

3. El conjunto $f^{-1}(\bar{o}) = \{\bar{u} \in V / f(\bar{u}) = \bar{o}\}$ es un subespacio vectorial de V, que se llama *núcleo* de la aplicación lineal f y se denotará por $\text{Nuc}(f)$.
4. Si el espacio V tiene dimensión finita, se verifica que:

$$\dim \text{Nuc}(f) + \dim \text{Im}(f) = \dim V$$

NOTA: Algunos autores, a la dimensión del núcleo la llaman «nulidad» de la aplicación lineal.

DEMOSTRACION

1. Debemos comprobar que, si $\bar{w}_1, \bar{w}_2 \in f(V)$ y para cualesquiera $\lambda_1, \lambda_2 \in K$, el vector $\lambda_1\bar{w}_1 + \lambda_2\bar{w}_2$ pertenece a $f(V)$. Así ocurre ya que $\bar{w}_1 = f(\bar{u}_1)$ y $\bar{w}_2 = f(\bar{u}_2)$ para unos ciertos $\bar{u}_1, \bar{u}_2 \in V$, luego

$$\lambda_1\bar{w}_1 + \lambda_2\bar{w}_2 = \lambda_1 f(\bar{u}_1) + \lambda_2 f(\bar{u}_2) = f(\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2)$$

que pertenece a $f(V)$ ya que $\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 \in V$.

2. Hay que comprobar que cualquier $\bar{w} \in f(V)$ es una combinación lineal de las $f(\bar{u}_i)$. Así ocurre ya que, como existe $\bar{u} \in V$ tal que $f(\bar{u}) = \bar{w}$ y como el sistema $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ genera V, resulta que existen ciertos escalares $\lambda_1, \lambda_2, ..., \lambda_p$ tales que

$$\bar{w} = f(\bar{u}) = f(\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p) = \lambda_1 f(\bar{u}_1) + \lambda_2 f(\bar{u}_2) + \cdots + \lambda_p f(\bar{u}_p)$$

3. Si $\bar{u}_1, \bar{u}_2 \in \text{Nuc}(f)$, hay que probar que $\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2$ también es del $\text{Nuc}(f)$, para cualesquiera $\lambda_1, \lambda_2 \in K$. Así es, ya que:

$$\bar{u}_1, \bar{u}_2 \in \text{Nuc}(f) \Rightarrow f(\bar{u}_1) = f(\bar{u}_2) = \bar{o} \Rightarrow$$
$$\Rightarrow f(\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2) = \lambda_1\bar{o} + \lambda_2\bar{o} = \bar{o} \Rightarrow \lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 \in \text{Nuc}(f)$$

4. Si V tiene dimensión finita, entonces también la tiene $\text{Nuc}(f)$; sea $B = (\bar{e}_1, ..., \bar{e}_p)$ una base de $\text{Nuc}(f)$. Completando esta base con cierto

sistema de vectores $S = (\bar{u}_1, ..., \bar{u}_q)$ se obtiene una base $B \cup S$ de V. Dado que $B \cup S$ es un sistema generador de V, resulta que el sistema

$$f(B \cup S) = (f(\bar{e}_1), ..., f(\bar{e}_p), f(\bar{u}_1), ..., f(\bar{u}_q))$$

es un generador de $f(V)$. Como $f(\bar{e}_i) = \bar{o}$ (para $i = 1, ..., p$) pues $\bar{e}_i$ es del núcleo, resulta que los q vectores

$$\bar{w}_1 = f(\bar{u}_1), ... \bar{w}_q = f(\bar{u}_q)$$

forman un sistema generador de $f(V)$. Si se demuestra que estos vectores son linealmente independientes, entonces formarían base de $f(V)$, que tendría entonces dimensión q, y por ello la propiedad quedaría comprobada, ya que la dimensión de Nuc (f) es p y que la de V es $p + q$.

Veamos, pues, que $(\bar{w}_1, ..., \bar{w}_q)$ es un sistema independiente. Para ello, supongamos que

$$\mu_1 \bar{w}_1 + \cdots + \mu_q \bar{w}_q = \bar{o} \quad \text{para ciertos } \mu_1, ..., \mu_q \in K$$

De ser así, como $\bar{w}_i = f(\bar{u}_i)$ y f es lineal, se obtiene que

$$\bar{o} = \mu_1 f(\bar{u}_1) + \cdots + \mu_q f(\bar{u}_q) = f(\mu_1 \bar{u}_1 + \cdots + \mu_q \bar{u}_q)$$

luego $\mu_1 \bar{u}_1 + \cdots + \mu_q \bar{u}_q$ pertenece al núcleo de f y, por ello, es combinación lineal de los vectores de la base B, es decir, para ciertos escalares λ_1, ..., λ_p se verificará que

$$\lambda_1 \bar{e}_1 + \cdots + \lambda_p \bar{e}_p + \mu_1 \bar{u}_1 + \cdots + \mu_q \bar{u}_q = \bar{o}$$

Ahora bien, como $B \cup S$ es un sistema independiente, de la última relación se deduce que todos sus escalares son nulos y en particular $\mu_1 = \cdots = \mu_q = 0$, de lo que se deduce que las $\bar{w}_1, ..., \bar{w}_q$ son linealmente independientes.

EJEMPLOS

1. Sean V y W los siguientes espacios reales:

$V = \{\text{funciones polinómicas } p(x) = a + bx + cx^2 + dx^3 \text{ (de grado } \leq 3)\}$

$W = \mathcal{F}(\mathbb{R}, \mathbb{R}) = \{\text{funciones de } \mathbb{R} \text{ en } \mathbb{R}\}$

Considérese la aplicación lineal $F: V \to W$ definida por

$$F(p(x)) = p'(x)$$

o sea

$$F(a + bx + cx^2 + dx^3) = b + 2cx + 3dx^2$$

La imagen de F es el subespacio de W formado por los polinomios de grado menor o igual que 2. El núcleo de F es el subespacio de V que forman los polinomios constantes. Nótese que:

$$\dim V = 4 \quad ; \quad \dim \mathrm{Im}(F) = 3 \quad ; \quad \dim \mathrm{Nuc}(F) = 1$$

2. Sea $f: \mathbb{R}^3 \to \mathbb{R}^4$ la aplicación lineal definida por

$$f(x, y, z) = (x + z, y - z, x + y, x - y + 2z)$$

Como $\vec{u}_1 = (1, 0, 0)$, $\vec{u}_2 = (0, 1, 0)$ y $\vec{u}_3 = (0, 0, 1)$ es un sistema generador de $\mathbb{R}^3$, los vectores

$$\vec{w}_1 = f(\vec{u}_1) = (1, 0, 1, 1) \quad ; \quad \vec{w}_2 = f(\vec{u}_2) = (0, 1, 1, -1) \quad ;$$
$$\vec{w}_3 = f(\vec{u}_3) = (1, -1, 0, 2)$$

generan $f(\mathbb{R}^3)$ y, por ser $\vec{w}_3 = \vec{w}_1 - \vec{w}_2$, resulta que

$$\mathrm{Im}(f) = \mathcal{V}(\vec{w}_1, \vec{w}_2) = \{(\alpha, \beta, \alpha + \beta, \alpha - \beta) / \alpha, \beta \in \mathbb{R}\} \subset \mathbb{R}^4$$

El núcleo de f está formado por la $(x, y, z) \in \mathbb{R}^3$ tales que

$$x + z = y - z = x + y = x - y + 2z = 0$$

Como este sistema de ecuaciones es equivalente al $x = -y = -z$, resulta:

$$\mathrm{Nuc}(f) = \{(x, -x, -x) / x \in \mathbb{R}\} \subset \mathbb{R}^3$$

Nótese que

$$\dim \mathbb{R}^3 = 3 \quad ; \quad \dim \mathrm{Im}(f) = 2 \quad ; \quad \dim \mathrm{Nuc}(f) = 1$$

EJERCICIO

Si $f: V \to W$ es una aplicación lineal entre espacios vectoriales (sobre un mismo cuerpo K), demuéstrese que:

a) Si V' es un subespacio de V, entonces $f(V')$ es subespacio de W.
b) Si W' es un subespacio de W, entonces $f^{-1}(W')$ es subespacio de V.

RESOLUCION

a) Esta cuestión se prueba de igual modo que la anterior propiedad 3.ª; basta con sustituir, allí, V por V'.
b) Hemos de comprobar que si $\vec{u}_1$ y $\vec{u}_2$ son de $f^{-1}(W')$, entonces también lo es el vector $\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2$ para cualesquiera escalares λ_1 y λ_2. Como

$\bar{w}_1 = f(\bar{u}_1)$ y $\bar{w}_2 = f(\bar{u}_2)$ son de W', resulta que el vector $\lambda_1\bar{w}_1 + \lambda_2\bar{w}_2$ también pertenece a este subespacio, es decir,

$$\lambda_1\bar{w}_1 + \lambda_2\bar{w}_2 = \lambda_1 f(\bar{u}_1) + \lambda_2 f(\bar{u}_2) = f(\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2) \in W'$$

lo que significa que $\lambda\bar{u} + \lambda\bar{u}_2$ es de $f^{-1}(W')$, como había que probar.

6.2. ISOMORFISMOS

APLICACIONES LINEALES INYECTIVAS

Según tendremos ocasión de ir comprobando más adelante, de entre las aplicaciones lineales, las inyectivas son especialmente importantes. Ellas son las que conservan la dimensión, las que tienen por núcleo al subespacio nulo, las que transforman una base cualquiera del espacio origen en una base del espacio imagen. Analicemos con detalle estas características de las aplicaciones lineales inyectivas:

[083]

Sea $f: V \to W$ una aplicaciión lineal entre espacios vectoriales, sobre el mismo cuerpo K. Se verifica que:

1. **La aplicación lineal f es inyectiva si y sólo si $\text{Nuc}(f) = O$.**
2. **Si V tiene dimensión finita, entonces f es inyectiva si y sólo si $\dim V = \dim f(V)$.**
3. **Si $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una base de V, entonces f es inyectiva si y sólo si $f(B) = (f(\bar{e}_1), f(\bar{e}_2), ..., f(\bar{e}_n))$ es una base de $f(V)$, es decir, si y sólo si $f(B)$ es un sistema independiente de vectores de W.**

DEMOSTRACION

1. Si f es inyectiva, de ser $f(\bar{u}) = \bar{o}$ para $\bar{u} = V$, como $f(\bar{o}) = \bar{o}$, los vectores $\bar{u}$ y $\bar{o}$ tendrían la misma imagen por f (inyectiva), luego sería $\bar{u} = \bar{o}$; así, pues, $\bar{o}$ es el único vector de $\text{Nuc}(f)$. Recíprocamente, si $\text{Nuc}(f) = O$, de ser $f(\bar{u}_1) = f(\bar{u}_2)$ para algunos $\bar{u}_1, \bar{u}_2 \in V$, sería

$$f(\bar{u}_1 - \bar{u}_2) = f(\bar{u}_1) - f(\bar{u}_2) = \bar{o}, \quad \text{luego } \bar{u}_1 - \bar{u}_2 \in \text{Nuc}(f)$$

y como $\text{Nuc}(f) = O$ resulta que $\bar{u}_1 - \bar{u}_2 = \bar{o}$, o sea $\bar{u}_1 = \bar{u}_2$, lo que prueba que f es inyectiva.

2. De acuerdo con la propiedad anterior y con la fórmula [082], 4) se obtiene:

$$f \text{ inyectiva} \Leftrightarrow \overline{\text{Nuc}}\,(f) = O \Leftrightarrow \dim \text{Nuc}\,(f) = 0 \Leftrightarrow$$
$$\Leftrightarrow \dim V = \dim f(V) + 0$$

3. Finalmente, dada una base $B = (\bar{e}_1, ..., \bar{e}_n)$ de V, como $f(B)$ es un sistema generador (de n vectores) de $f(V)$, se verifica que:

$$f(B) \text{ base de } f(V) \Leftrightarrow f(B) \text{ independientemente} \Leftrightarrow \text{rang } f(B) = n \Leftrightarrow$$
$$\Leftrightarrow \dim f(V) = n = \dim V \Leftrightarrow f \text{ inyectiva}$$

OBSERVACION

Una aplicación lineal $f: V \to W$, entre espacios vectoriales sobre un mismo cuerpo K, es inyectiva si y sólo si todo sistema linealmente independiente de vectores de V tiene por imagen a un sistema de vectores de W que también es linealmente independiente.

COMPROBACION

- Si f es inyectiva, dado un sistema $(\bar{u}_1, ... \bar{u}_p)$ de vectores independientes de V, el sistema imagen $(f(\bar{u}_1), ..., f(\bar{u}_p))$ también es linealmente independiente, puesto que (para $\lambda_1, ..., \lambda_p \in K$):

$$\sum_{i=1}^{p} \lambda_i f(\bar{u}_i) = \bar{o} \Rightarrow f\left(\sum_{i=1}^{p} \lambda_i \bar{u}_i\right) = \bar{o} \Rightarrow \sum_{i=1}^{p} \lambda_i \bar{u}_i = \bar{o} \Rightarrow$$
$$\Rightarrow \lambda_1 = \cdots = \lambda_p = 0$$

- Si la imagen de cualquier sistema independiente de V es un sistema independiente de W, se verificará en particular que dado $\bar{u} \neq \bar{o}$ en V, ha de ser $f(\bar{u}) \neq \bar{o}$. De ello resulta, pues, que $\text{Nuc}\,(f) = O$ y por tanto f es inyectiva.

EJEMPLO

La aplicación lineal $f: \mathbb{R}^3 \to \mathbb{R}^4$ definida mediante

$$f(x, y, z) = (x, x + y, y + z, x + y + z)$$

es inyectiva ya que las imágenes de los vectores de la base canónica $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ de $\mathbb{R}^3$ son los vectores

$$f(1, 0, 0) = (1, 1, 0, 1) \quad ; \quad f(0, 1, 0) = (0, 1, 1, 1) \quad ; \quad f(0, 0, 1) = (0, 0, 1, 1)$$

que forman una base de la imagen, ya que son linealmente independientes, pues:

$$\operatorname{rang}(f(\bar{e}_1), f(\bar{e}_2), f(\bar{e}_3)) = \operatorname{rang}\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = 3$$

La inyectividad de f también se podía haber comprobado viendo que $\operatorname{Nuc}(f) = O$; así es, ya que:

$$\operatorname{Nuc}(f): \begin{cases} x & = 0 \\ x+y & = 0 \\ y+z & = 0 \\ x+y+z & = 0 \end{cases} \Leftrightarrow \operatorname{Nuc}(f): \begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases} \Leftrightarrow \operatorname{Nuc}(f) = O$$

ISOMORFISMOS

Cuando se dispone de una aplicación lineal, entre dos espacios vectoriales V y W, que además es biyectiva, puede considerarse que V y W son iguales, se pueden identificar. Desde el punto de vista de los espacios vectoriales, no hay nada que permita diferenciar a V de W. Estas aplicaciones lineales y biyectivas se llaman isomorfismos:

[084]

Se llama *isomorfismo* a una aplicación $f: V \to W$, entre espacios vectoriales sobre el mismo cuerpo K, que sea lineal y biyectiva. Si $f: V \to W$ es un isomorfismo, los dos espacios vectoriales se dicen *isomorfos*. Un *isomorfismo* $f: V \to V$, de un espacio en sí mismo, recibe el nombre de *automorfismo*.

1. La composición de dos isomorfismos es, también, un isomorfismo.
2. Una aplicación lineal $f: V \to W$ es isomorfismo si y sólo si:
$$\operatorname{Im}(f) = W \quad \text{y} \quad \operatorname{Nuc}(f) = O$$
3. Si V tiene dimensión finita, una aplicación lineal $f: V \to W$ es un isomorfismo si y sólo si:
$$\dim V = \dim f(V) = \dim W$$
4. Si V tiene dimensión finita, una aplicación lineal $f: V \to V$ es automorfimo si y sólo si es inyectiva o si y sólo si es sobreyectiva.
5. Si $f: V \to W$ es un isomorfismo, entonces $f^{-1}: W \to V$ también es un isomorfismo.
6. Dos espacios vectoriales de dimensión finita (sobre el mismo cuerpo K) son isomorfos si y sólo si tienen la misma dimensión.

DEMOSTRACION

1. Como al componer aplicaciones lineales se obtienen aplicaciones lineales y al componer biyecciones se obtienen biyecciones, resulta obvio que al componer isomorfismos se obtienen isomorfismos.
2. Por definición, f es sobreyectiva si y sólo si $\text{Im}(f) = W$. Según [083], 1), f es inyectiva si y sólo si $\text{Nuc}(f) = O$.
3. Esta propiedad se deduce fácilmente de:

$$f \text{ sobreyectiva} \Leftrightarrow f(V) = W \Leftrightarrow \dim f(V) = \dim W$$

$$f \text{ inyectiva} \Leftrightarrow \dim \text{Nuc}(f) = 0 \Leftrightarrow \dim V = \dim f(V)$$

4. Si en el caso anterior se toma $W = V$, resulta que f es isomorfismo si y sólo si $\dim V = \dim f(V)$, es decir, si y sólo si f es sobreyectiva. La anterior igualdad entre dimensiones equivale a la $\dim \text{Nuc}(f) = 0$, o sea, a que f sea inyectiva.
5. Nótese primeramente que, como $f: V \to W$ es biyectiva, existe su aplicación recíproca $f^{-1}: W \to V$, que es biyectiva. Hay que probar que f^{-1} es lineal, es decir, que $f^{-1}(\lambda\bar{u} + \mu\bar{v})$ es igual a $\lambda f^{-1}(\bar{u}) + \mu f^{-1}(\bar{v})$ para cualesquiera $\bar{u}, \bar{v} \in W$ y $\lambda, \mu \in K$; dicha igualdad equivale a la igualdad de sus imágenes por f y esta última es cierta ya que, como:

$$f(f^{-1}(\lambda\bar{u} + \mu\bar{v})) = \lambda\bar{u} + \mu\bar{v}$$

y, dado que f es lineal, es

$$f(\lambda f^{-1}(\bar{u}) + \mu f^{-1}(\bar{v})) = \lambda f(f^{-1}(\bar{u})) + \mu f(f^{-1}(\bar{v})) = \lambda\bar{u} + \mu\bar{v}$$

6. Según se probó en [068], si $\dim V = n$, el espacio V es isomorfo al K^n; igualmente, si $\dim W = n$, también W es isomorfo a K^n. Si V y W tienen la misma dimensión, n, existen pues dos isomorfismos $f: V \to K^n$ y $g: K^n \to W$, luego V y W son isomorfos ya que $g \circ f: V \to W$ es un isomorfismo. Recíprocamente, si V y W son isomorfos, entonces existe un isomorfismo $f: V \to W$ y, según la anterior propiedad 3.ª, ha de ser $\dim V = \dim W$.

EJEMPLOS

1. Sea V un espacio vectorial, sea U un subespacio de V y sean U_1 y U_2 dos subespacios suplementarios de U respecto de V, esto es, tales que $V = U \oplus U_1$ y $V = U \oplus U_2$. Para cualquier $\bar{u}_1 \in U_1$, como entonces $\bar{u}_1 \in V$, existen unos únicos $\bar{u} \in U$ y $\bar{u}_2 \in U_2$ tales que $\bar{u}_1 = \bar{u} + \bar{u}_2$; pues bien, la aplicación $f: U_1 \to U_2$ definida mediante $f(\bar{u}_1) = \bar{u}_2$ es un isomorfismo de U_1 en U_2.

2. El espacio vectorial $\mathbb{R}^n$ es isomorfo al espacio vectorial V de los polinomios de grado menor que n. Un isomorfismo entre ellos lo es la aplicación $f: V \to \mathbb{R}^n$ dada por

$$f(a_0 + a_1x + \cdots + a_{n-1}x^{n-1}) = (a_0, a_1, \ldots, a_{n-1})$$

β MATRICES DE LAS APLICACIONES LINEALES

En el caso de dimensión finita, para manejarnos debidamente con las aplicaciones lineales, vamos a recurrir a las coordenadas. Es decir, dada una aplicación lineal $f: V \to W$ y elegidas sendas bases $(\bar{e}_i)$ en V y $(\bar{u}_j)$ en W, nos vamos a ocupar en obtener las coordenadas y_j, de la imagen $\bar{y} = f(\bar{x})$ en W, en función de las coordenadas x_i, del vector origen $\bar{x}$ en V. En las ecuaciones que proporcionan las y_j, en función de las x_i, intervienen unos coeficientes que caracterizan a la aplicación f y son las coordenadas, en la base $(\bar{u}_j)$ de W, de los vectores $f(\bar{e}_i)$, imágenes de los vectores de la base $(\bar{e}_i)$ de V. Dichos coeficientes intervienen en las ecuaciones con estructura de matriz; en ésta, las columnas son las coordenadas de los vectores $f(\bar{e}_i)$.

Hemos de señalar que, si V tiene dimensión finita n, la aplicación lineal $f: V \to W$ queda definida de modo inequívoco en cuanto se conozcan las n imágenes $f(\bar{e}_i)$, de los vectores de una base $(\bar{e}_i)$ de V. Con sólo conocer estos n vectores imagen, y por el hecho de ser f lineal, quedan determinadas las imágenes de todos los vectores de V.

6.3. ECUACIONES Y MATRIZ DE UNA APLICACION LINEAL

❑ DETERMINACION DE UNA APLICACION LINEAL

[085]

Sea V y W dos espacios vectoriales, sobre un mismo cuerpo K. Si $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una base de V y dado un sistema cualquiera $S = (\bar{w}_1, \bar{w}_2, ..., \bar{w}_n)$ de vectores de W, existe una aplicación lineal $f: V \to W$, y una sola, tal que

$$f(\bar{e}_1) = \bar{w}_1,\ f(\bar{e}_2) = \bar{w}_2,\ ...,\ f(\bar{e}_n) = \bar{w}_n$$

DEMOSTRACION

Vamos a recurrir a las coordenadas de los vectores de V en la base B; para cada vector $\bar{x} \in V$, llamaremos $(x_1, x_2, ..., x_n)$ a su sistema de coordenadas. De

existir una tal aplicación lineal f, habría de verificarse que:

$$f(\bar{x}) = f\left(\sum_{i=1}^{n} x_i \bar{e}_i\right) = \sum_{i=1}^{n} x_i f(\bar{e}_i) = \sum_{i=1}^{n} x_i \bar{w}_i$$

Así pues, la única solución posible del problema es la aplicación

$$\bar{x} \rightarrowtail f(\bar{x}) = x_1 \bar{w}_1 + x_2 \bar{w}_2 + \cdots + x_n \bar{w}_n$$

Si esta aplicación fuera lineal, la demostración habría acabado. Veamos que así es: para cualesquiera $\bar{x}, \bar{x}' \in V$ y $\lambda, \lambda' \in K$, se tiene:

$$f(\lambda\bar{x} + \lambda'\bar{x}') = \sum_{i=1}^{n} (\text{coordenada } i\text{-ésima de } \lambda\bar{x} + \lambda'\bar{x}')\bar{w}_i =$$

$$= \sum_{i=1}^{n} (\lambda x_i + \lambda' x_i')\bar{w}_i = \lambda \sum_{i=1}^{n} x_i \bar{w}_i + \lambda' \sum_{i=1}^{n} x_i' \bar{w}_i = \lambda f(\bar{x}) + \lambda' f(\bar{x}')$$

EJEMPLO

Sea $f : \mathbb{R}^3 \rightarrow \mathcal{M}_{2\times 2}$ una aplicación lineal de la que se sabe que

$$f(1, 0, 0) = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} \quad ; \quad f(0, 1, 0) = \begin{bmatrix} 1 & -1 \\ -5 & 5 \end{bmatrix}$$

$$f(0, 0, 1) = \begin{bmatrix} 2 & -2 \\ -3 & 4 \end{bmatrix}$$

Como los vectores $\bar{e}_1 = (1, 0, 0)$, $\bar{e}_2 = (0, 1, 0)$ y $\bar{e}_3 = (0, 0, 1)$ forman una base de $\mathbb{R}^3$ (la base canónica), se puede asegurar que hay una y sólo una aplicación f que cumple lo exigido; dicha aplicación está dada por:

$$f(x, y, z) = xf(1, 0, 0) + yf(0, 1, 0) + zf(0, 0, 1) =$$

$$= \begin{bmatrix} 3x + y + 2z & x - y - 2z \\ 2x - 5y - 3z & 4x + 5y + 4z \end{bmatrix}$$

ECUACIONES DE UNA APLICACION LINEAL

[086]

Sean V y W dos espacios vectoriales, sobre el mismo cuerpo K. Sea $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ una base de V en la que las coordenadas de $\bar{x} \in V$ son $(x_1, x_2, ..., x_n)$; sea $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_m)$ una base de W en la que las coordenadas de $\bar{y} \in W$ son $(y_1, y_2, ..., y_m)$. Si $f: V \to W$ es la aplicación lineal que transforma $\bar{e}_j$ en $f(\bar{e}_j) = \bar{w}_j$ (para $j = 1, 2, ..., n$) y si, utilizando coordenadas, estos vectores son los $\bar{w}_j = a_{1j}\bar{u}_1 + a_{2j}\bar{u}_2 + \cdots + a_{mj}\bar{u}_m$, entonces la imagen $\bar{y} = f(\bar{x})$ de un $\bar{x} \in V$ viene dada, por sus coordenadas, mediante $y_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$ (para $i = 1, 2, ..., m$). Resumiendo:

$$\begin{array}{l} f(\bar{e}_1) = a_{11}\bar{u}_1 + a_{21}\bar{u}_2 + \cdots + a_{m1}\bar{u}_m \\ f(\bar{e}_2) = a_{12}\bar{u}_1 + a_{22}\bar{u}_2 + \cdots + a_{m2}\bar{u}_m \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f(\bar{e}_n) = a_{1n}\bar{u}_1 + a_{2n}\bar{u}_2 + \cdots + a_{mn}\bar{u}_m \end{array} \Rightarrow \begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases}$$

$$\left(\begin{array}{c} f(\bar{e}_j) = \sum_{i=1}^{m} a_{ij}\bar{u}_i \\ \text{para } j = 1, 2, ..., n \end{array} \right) \Rightarrow \left(\begin{array}{c} y_i = \sum_{j=1}^{n} a_{ij}x_j \\ \text{para } i = 1, 2, ..., m \end{array} \right)$$

COMPROBACION

Como f es lineal y recurriendo a las coordenadas de $\bar{x} \in V$, se obtiene (para $j = 1, 2, ..., n$ e $i = 1, 2, ..., m$):

$$\bar{y} = f(\bar{x}) = f(\sum_j x_j \bar{e}_j) = \sum_j x_j f(\bar{e}_j) = \sum_j x_j (\sum_i a_{ij}\bar{u}_i) =$$

$$= \sum_j (\sum_i x_j a_{ij}\bar{u}_i) = \sum_i (\sum_j a_{ij}x_j)\bar{u}_i$$

De esta relación se deduce que $\sum_j a_{ij}x_j$ es la coordenada i-ésima de $\bar{y}$, es decir, $y_i = \sum_j a_{ij}x_j$.

EJEMPLO

Consideremos la aplicación lineal $f: \mathbb{R}^3 \to \mathbb{R}^4$ en la que

$$f(1, 0, 0) = (3, 2, -1, 1) \quad ; \quad f(1, 1, 0) = (5, -4, 1, -2)$$

$$f(1, 1, 1) = (2, 1, -6, 3)$$

Recurriendo a la linealidad de f se obtienen fácilmente las imágenes de los

vectores de la base canónica de $\mathbb{R}^3$:

$$f(1, 0, 0) = (3, 2, -1, 1)$$
$$f(0, 1, 0) = f(1, 1, 0) - f(1, 0, 0) = (2, -6, 2, -3)$$
$$f(0, 0, 1) = f(1, 1, 1) - f(1, 1, 0) = (-3, 5, -7, 5)$$

Recurriendo ya al resultado anterior, se obtiene que

$$f(x_1, x_2, x_3) = (y_1, y_2, y_3, y_4)$$

viene dado por:

$$\begin{cases} y_1 = 3x_1 + 2x_2 - 3x_3 \\ y_2 = 2x_1 - 6x_2 + 5x_3 \\ y_3 = -1x_1 + 2x_2 - 7x_3 \\ y_4 = x_1 - 3x_2 + 5x_3 \end{cases}$$

MATRIZ DE UNA APLICACION LINEAL

De acuerdo con el anterior resultado, en el caso de dimensión finita una aplicación lineal $\bar{x} \rightarrowtail \bar{y} = f(\bar{x})$ queda determinada por unos ciertos escalares a_{ij} (coeficiente de x_i en la expresión que da y_j). Estos escalares intervienen, en la ecuación de f, formando una matriz, de la que vamos a ocuparnos. En términos de matrices, las conclusiones que se acaban de obtener acerca de la ecuación de una aplicación lineal se pueden expresar como sigue:

[087]

Sean V y W espacios vectoriales de dimensión finita, sobre un mismo cuerpo K. Sea $(\bar{e}_1, ..., \bar{e}_n)$ una base de V y sea $(\bar{u}_1, ..., \bar{u}_m)$ una base de W; en ellas, llamaremos $(x_1, ..., x_n)$ e $(y_1, ..., y_m)$ a las coordenadas de $\bar{x} \in V$ y de $\bar{y} \in W$.

Si $f: V \rightarrow W$ es la aplicación lineal en la que $f(\bar{e}_j) = \sum_{i=1}^{m} a_{ij}\bar{u}_i$ (para $j = 1, 2, ..., n$), entonces $\bar{x} \rightarrowtail \bar{y} = f(\bar{x})$ admite la siguiente ecuación matricial, respecto de las bases anteriores:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

o abreviadamente

$$[y_i] = [a_{ij}][x_j]$$

ó $Y = AX$

Se dice que $A = [a_{ij}]$ es la *matriz de la aplicación lineal* f respecto de las bases $(\bar{e}_j)$ de V y $(\bar{u}_i)$ de W.

NOTA: Como se habrá observado, las columnas de A son las matrices columna de las coordenadas, en la base $(\bar{u}_i)$, de las imágenes $f(\bar{e}_j)$, de los vectores de la base $(\bar{e}_j)$.

EJERCICIO

Hallar el núcleo y la imagen de la aplicación lineal $f: \mathbb{R}^4 \to \mathbb{R}^3$ tal que:

$$f(1, 0, 0, 0) = (1, -1, 2) \quad ; \quad f(0.\ 0, 1, 0) = (4, -1, 5)$$
$$f(0, 1, 0, 0) = (2, 1, 1) \quad ; \quad f(0.\ 0, 0, 1) = (-1, -5, 4)$$

En concreto, se piden sendas bases del núcleo y de la imagen.

RESOLUCION

Respecto de las bases canónicas de $\mathbb{R}^4$ y $\mathbb{R}^3$, la ecuación matricial, $Y = AX$, de la aplicación lineal es:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 & -1 \\ -1 & 1 & -1 & -5 \\ 2 & 1 & 5 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

- El núcleo lo forman las soluciones de $f(\bar{x}) = \bar{o}$, que si se usan coordenadas se pone en la forma $AX = O$. Para resolver este sistema lineal homogéneo, realizando operaciones elementales en las filas de su matriz A, se obtienen sucesivamente las matrices:

$$\begin{bmatrix} 1 & 2 & 4 & -1 \\ 0 & 3 & 3 & -6 \\ 0 & -3 & -3 & 6 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 4 & -1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Por tanto, los vectores del núcleo son los que cumplen (para $\alpha, \beta \in \mathbb{R}$):

$$\left.\begin{aligned} x_1 &= -2\alpha - 3\beta \\ x_2 &= -\alpha + 2\beta \\ x_3 &= \alpha \\ x_4 &= \beta \end{aligned}\right\} \text{ luego}$$

$(-2, -1, 1, 0)$ y $(-3, 2, 0, 1)$ forman base de $\mathrm{Nuc}(f)$

- La imagen es el subespacio de $\mathbb{R}^3$ que engendran las cuatro columnas de A. Realizando operaciones elementales en las columnas de A se obtienen sistemas de columnas equivalentes al dado; al proceder de este modo, se obtiene sucesivamente:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 3 & 3 & -6 \\ 2 & -3 & -3 & 6 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

Por tanto, los dos vectores columna no nulos de la última matriz generan la imagen y, como son independientes, son base de ella; esto es:

$$(1, 1, 0) \text{ y } (0, 1, -1) \quad \text{forman base de } \mathrm{Im}(f)$$

OBSERVACION

[088]

Sean V_n y W_m espacios vectoriales de dimensiones n y m, ambos sobre un mismo cuerpo K, en los que se consideran sendas bases (por ejemplo, $V_n = K^n$ y $W_m = K^m$ con las bases canónicas). Al conjunto de las aplicaciones lineales de V_n en W_m se le denotará por $\mathcal{L}(V_n, W_m)$; se pone $\mathcal{M}_{m\times n}$ para designar al conjunto de las matrices de tamaño $m \times n$ (con elementos del cuerpo K). La aplicación

$$\mathcal{L}(V_n, W_m) \to \mathcal{M}_{m\times n} \quad , \quad f \mapsto A = \text{matriz de } f$$

que a cada aplicación lineal le hace corresponder su matriz asociada en las bases elegidas, es una biyección.

Por ello y abusando del lenguaje, hay ocasiones en las que se confunde una aplicación lineal con su matriz asociada. Así, por ejemplo, hay quien dice «sea la aplicación lineal A» en lugar de decir «sea la aplicación lineal cuya matriz es A».

RANGO DE LA MATRIZ ASOCIADA

En el caso de dimensión finita, una aplicación lineal y su matriz asociada, respecto de cualesquiera bases, tienen el mismo rango. En particular, las matrices asociadas a los isomorfismos son las matrices regulares. Es decir:

[089]

Sea $f: V \to W$ una aplicación lineal entre espacios vectoriales de dimensión finita, ambos sobre el mismo cuerpo K, y sea A la matriz asociada a f, respecto de cualesquiera bases de V y W. Se verifica que:

1. El rango de la aplicación lineal f es igual al rango de su matriz asociada A.
2. La aplicación lineal f es un isomorfismo(*) si y sólo si su matriz asociada A es regular.

(*) Nótese que aquí, en este apartado 2, se supone que $\dim V = \dim W$ y, por tanto, A es cuadrada.

COMPROBACION

1. Llamemos $(\bar{e}_1, ..., \bar{e}_n)$ a la base que se toma en V, con lo que $S = (f(\bar{e}_1), ..., f(\bar{e}_n))$ es un sistema generador de $f(V)$. Recuérdese que las columnas de A son las columnas de coordenadas de los vectores de S. De acuerdo con la definición del rango de una aplicación lineal (véase [082]), de lo que se acaba de decir se desprende que:

$$\text{rang } f = \text{rang } f(V) = \text{rang } S = \text{rang de columnas de } A = \text{rang } A$$

2. En este caso, se consideran espacios de la misma dimensión, $\dim V = \dim W = n$, con lo que A es cuadrada, de tamaño $n \times n$. Nótese que f es isomorfismo si y sólo si $\text{rang } f = n$; nótese también que A es regular si y sólo si $\text{rang } A = n$. Resulta entonces obvio que f es inyectiva si y sólo si A es regular.

EJEMPLO

Sea V el espacio vectorial de los polinomios reales de grado menor o igual que 2 y considérese la aplicación lineal $f: V \to \mathbb{R}^3$ definida por

$$f(a + bx + cx^2) = (a + b - 2c,\ 2a + b + c,\ 2a + 3b + hc)$$

Hallar $h \in \mathbb{R}$ para que f no sea un isomorfismo. Para dicho valor de h, hallar una base de la imagen $\text{Im}(f)$.

RESOLUCION

El valor buscado de h es aquel que hace singular a la matriz A de f (en la base usual de V y en la canónica de $\mathbb{R}^3$); realizando, pues, operaciones elementales en las columnas de A, se tiene sucesivamente:

$$A = \begin{bmatrix} 1 & 1 & -2 \\ 2 & 1 & 1 \\ 2 & 3 & h \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 5 \\ 2 & 1 & h+4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -1 & h+9 \end{bmatrix}$$

La matriz A es singular para $h = -9$. Para este valor, el rango de A vale 2, la dimensión de $\text{Im}(f)$ es 2 y una base de este espacio la forman las dos columnas no nulas de la última matriz, es decir, los vectores (1, 2, 2) y (0, 1, −1) de $\mathbb{R}^3$.

6.4. MATRICES EQUIVALENTES

MATRICES ASOCIADAS A UN MISMO HOMOMORFISMO

Cuando (en [087]) hablamos de la matriz A asociada a una aplicación lineal $f: V \to W$, se eligieron sendas bases en V y W y, respecto de ellas, se determinó dicha matriz A. Si las bases de V y W se cambian por otras, la matriz asociada será distinta de la anterior. Nos ocupamos aquí de analizar la relación que liga a la nueva matriz asociada, A', con la antigua A, en función de las matrices de los cambios de coordenadas que se hagan en V y W. El resultado de este análisis será que A y A' son matrices equivalentes (véase [020]):

[090]

Sea $f: V \to W$ una aplicación lineal entre espacios vectoriales de dimensión finita, ambos sobre un mismo cuerpo K, y sea A la matriz asociada a f respecto de ciertas bases de V y W. Si en V y W se cambian las bases y llamando P y Q a las matrices de los respectivos cambios de coordenadas, entonces la matriz asociada a f en las nuevas bases es la $A' = Q^{-1}AP$.

Dos matrices A y A', ambas de igual tamaño $m \times n$, se dicen *matrices equivalentes*(*) si están asociadas a una misma aplicación lineal de K^n en K^m (respecto de bases adecuadas) o, lo que es lo mismo, si existen dos matrices regulares P y Q, de tamaños respectivos $n \times n$ y $m \times m$, tales que $A' = Q^{-1}AP$.

(*) En [020] se dijo que dos matrices $m \times n$ eran equivalentes si tenían el mismo rango; enseguida veremos que ésta y aquélla definición coinciden.

DEMOSTRACION

Llamemos X e Y a las matrices columna de coordenadas de un $\bar{x} \in V$ y de $\bar{y} = f(\bar{x})$, respecto de las bases primitivas. Entonces la ecuación matricial de f en dichas bases es $Y = AX$. Si X' e Y' son las matrices columna de coordenadas de $\bar{x}$ e $\bar{y}$ en las bases nuevas, la ecuación de f en dichas bases es $Y' = A'X'$ (para una cierta matriz A', que buscamos). Como sabemos que $X = PX'$ y que $Y = QY'$, se puede poner:

$$\left.\begin{array}{l} Y = AX \\ X = PX' \\ Y = QY' \end{array}\right] \Rightarrow QY' = APX' \Rightarrow \left.\begin{array}{l} Y' = (Q^{-1}AP)X' \\ Y' = A'X' \end{array}\right]^{(*)} \Rightarrow A' = Q^{-1}AP$$

(*) Estas dos relaciones se verifican para toda matriz columna de coordenadas X'.

La implicación recíproca afirma que si $A' = Q^{-1}AP$, donde P y Q son regulares, entonces A y A' están asociadas a una misma aplicación lineal (en unas ciertas bases). Esto se comprueba fácilmente:

Sea $f: K^n \to K^m$ la aplicación lineal que respecto de las bases canónicas (por ejemplo) tiene por ecuación a la $Y = AX$. Realizando, en $\mathbb{R}^n$ y en $\mathbb{R}^m$, los cambios de coordenadas $X = PX'$ e $Y = QY'$, se pasa a operar en unas nuevas bases en las que la matriz de f es la $Q^{-1}AP$ (según se acaba de ver); es decir, es A'.

EJEMPLO

Sea $f: \mathbb{R}^3 \to \mathbb{R}^2$ la aplicación lineal que, respecto de las bases canónicas $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ de $\mathbb{R}^3$ y $(\bar{u}_1, \bar{u}_2)$ de $\mathbb{R}^2$, tiene por ecuación a:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & -2 \\ -1 & 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad ; \quad \begin{bmatrix} 3 & 0 & -2 \\ -1 & 4 & 5 \end{bmatrix} \text{ es la matriz de } f$$

Al cambiar de bases en $\mathbb{R}^3$ y $\mathbb{R}^3$, pasando a las nuevas bases $(\bar{e}'_1, \bar{e}'_2, \bar{e}'_3)$ y $(\bar{u}'_1, \bar{u}'_2)$:

$$\begin{cases} \bar{e}'_1 = (1, 3, 0) = \bar{e}_1 + 3\bar{e}_2 \\ \bar{e}'_2 = (1, 0, 2) = \bar{e}_1 + 2\bar{e}_3 \\ \bar{e}'_3 = (0, 4, -2) = 4\bar{e}_2 - 2\bar{e}_3 \end{cases} \qquad \begin{cases} \bar{u}'_1 = (2, 1) = 2\bar{u}_1 + \bar{u}_2 \\ \bar{u}'_2 = (4, 3) = 4\bar{u}_1 + 3\bar{u}_2 \end{cases}$$

la matriz de f pasa a ser

$$A' = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 3 & 0 & -2 \\ -1 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 3 & 0 & 4 \\ 0 & 2 & -2 \end{bmatrix} = \begin{bmatrix} 4 & -6 & 3 \\ 5 & 11 & -2 \end{bmatrix}$$

❑ PROPOSICION

[091]

La equivalencia entre matrices de igual tamaño es una relación reflexiva, simétrica y transitiva (esto es, es una de las llamadas relaciones de equivalencia(*).

(*) Sobre las relaciones de equivalencia, véase [368], en el Apéndice A.4, al final del libro.

DEMOSTRACION

La equivalencia entre matrices es reflexiva (toda matriz A es equivalente a sí misma), ya que $A = I_m^{-1}AI_n$, donde I_m e I_n son matrices unidad (de los pertinentes tamaños), que son regulares. Es simétrica (si A es equivalente a A', entonces A' es equivalente a A), ya que si es $A' = Q^{-1}AP$ entonces es $A = QA'P^{-1}$, que se puede poner $A = S^{-1}A'R$, donde $R = P^{-1}$ y $S = Q^{-1}$ son regulares. Es transitiva (si A es equivalente a A' y A' es equivalente a A'', entonces A es equivalente a A''), ya que si $A' = Q^{-1}AP$ (con P y Q regulares) y $A'' = S^{-1}A'R$ (con R y S regulares), entonces

$$A'' = S^{-1}Q^{-1}APR = (QS)^{-1}A(PR)$$

donde PR y QS son matrices regulares.

MATRICES EQUIVALENTES Y RANGO

Según ya se ha hecho notar hace poco, hemos dado dos definiciones distintas de equivalencia: Primeramente, en [020], se dijo que dos matrices de igual tamaño eran equivalentes si tenían el mismo rango; en [090] acabamos de decir que las matrices equivalentes son las que están asociadas a una misma aplicación lineal. Nos ocuparemos aquí en comprobar que ambas definiciones son equivalentes.

Vamos también a relacionar todo ello con algo de lo que ya se habló anteriormente en [020] y en [026]; nos estamos refiriendo a las matrices canónicas de equivalencia.

[092]

Sea dada una matriz A de tamaño $m \times n$ (cuyos elementos pertenecen a un cuerpo K) que tiene rango r. Llamemos C_r a la siguiente matriz de tamaño $m \times n$:

$$C_r = \left[\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right]$$

donde

$$\begin{cases} I_r = \text{matriz de unidad de tamaño } r \times r \\ O \text{ denota a matrices nulas} \end{cases}$$

Se verifica que las matrices A y C_r están asociadas a una misma aplicación lineal (de K^n en K^m), respecto de bases adecuadas; es decir, A y C_r son equivalentes. Este resultado se expresa diciendo que A es *diagonalizable por equivalencia* y que C_r es la matriz que resulta de hacer la diagonalización. Se dice que C_r es la *matriz canónica de equivalencia* de A.

DEMOSTRACION

Sea $f: K^n \to K^m$ la aplicación lineal que en ciertas bases (las canónicas, por ejemplo) tiene a A por matriz asociada. Sea $(\bar{e}_1, ..., \bar{e}_s)$ una base de $\text{Nuc}(f)$, que completaremos con ciertos vectores $(\bar{u}_1, ..., \bar{u}_r)$ hasta obtener una base de K^n; nótese que $s + r = n$. Los vectores $\bar{w}_1 = f(\bar{u}_1), ..., \bar{w}_r = f(\bar{u}_r)$ forman una base de $\text{Im}(f)$; completemos adecuadamente esta base, con ciertos vectores $\bar{w}_{r+1}, ..., \bar{w}_m$, hasta obtener una base de K^m. La matriz asociada a f respecto de las bases

$$B_n = (\bar{u}_1, ..., \bar{u}_r, \bar{e}_1, ..., \bar{e}_s) \text{ de } K^n$$
$$B_m = (\bar{w}_1, ..., \bar{w}_r, \bar{w}_{r+1}, ..., \bar{w}_m) \text{ de } K^m$$

no es otra que C_r, ya que sus columnas son las matrices columna de coordenadas, respecto de la base B_m, de las imágenes de los vectores de B_n, que son $\bar{w}_1, ..., \bar{w}_r, \bar{0}, ..., \bar{0}$.

[093]

De lo hasta aquí dicho(*), se deduce que dos matrices A y A' son equivalentes, si y sólo si se verifica una cualquiera de las tres condiciones siguientes (las cuales son equivalentes entre sí):

1. Están asociadas a una misma aplicación lineal, respecto de bases adecuadas.
2. Las matrices A y A' son del mismo tamaño y tales que $A' = Q^{-1}AP$ para ciertas matrices cuadradas regulares P y Q.
3. Las matrices A y A' son del mismo tamaño y tienen igual rango.

(*) Véase [020] y [090].

DEMOSTRACION

La equivalencia entre 1 y 2 ya se comprobó en [090].

La condición 1 implica la 3, ya que el rango de una matriz es igual al rango de su aplicación lineal asociada (véase [089], 1) y, entonces, como A y A' están asociadas a una misma aplicación lineal, tienen ambas el mismo rango que ella (además de tener, obviamente, el mismo tamaño).

La condición 3 implica la 2. En efecto: la 3 implica que A y A' tengan, ambas, la misma matriz canónica de equivalencia C_r (llamamos r al rango de A y de A') y que, por ello, se pueda poner

$$C_r = Q^{-1}AP \quad \text{y} \quad C_r = Q'^{-1}A'P'$$

para ciertas matrices regulares P, P', Q y Q', de donde:

$$A' = Q'C_rP'^{-1} = Q^{-1}Q'APP'^{-1} = (Q'^{-1}Q)^{-1}A(PP'^{-1})$$

es decir, se cumple la condición 2, ya que $Q'^{-1}Q$ y PP'^{-1} son regulares.

OBSERVACIONES

Vamos a hacer algunas consideraciones acerca de la relación

$$Q^{-1}AP = C_r$$

que guarda una matriz A, de tamaño $m \times n$ y rango r, con su matriz canónica de equivalencia C_r.

1. Recuérdese que (véase [026]) existen dos operaciones elementales e' y e'' tales que aplicando e' a las filas de A y aplicando, después, e'' a las columnas de la matriz ya obtenida, se llega finalmente a la matriz C_r. Si e' se aplica también a las filas de la matriz unidad $m \times m$, se obtiene la matriz Q^{-1}; si e'' se aplica también a las columnas de la matriz unidad $n \times n$, se obtiene la matriz P, esto es:
   En $C_r = Q^{-1}AP$, las matrices C_r, P y Q son:

$$C_r = e_c''[e_f'(A)] \quad ; \quad Q^{-1} = e_f'(I_m) \quad ; \quad P = e_c''(I_n)$$

2. En la reciente demostración de diagonalizabilidad por equivalencia de la matriz A, se han hecho unos razonamientos, relacionados con la imagen y el núcleo de la aplicación lineal asociada a A, que nos permiten obtener las siguientes conclusiones:
   Si $f: V \to W$ es la aplicación lineal que, respecto de ciertas bases, tiene a A como matriz asociada (de tamaño $m \times n$ y rango r) y si $C_r = Q^{-1}AP$ es su matriz canónica de equivalencia, entonces:

   - las r primeras columnas de P son las columnas de coordenadas de los vectores de una base de $I_m(f)$,
   - las $n - r$ últimas columnas de Q son las columnas de coordenadas de los vectores de una base de $\mathrm{Nuc}\,(f)$.

3. Consideremos ahora el conjunto de todas las matrices de un cierto tamaño $m \times n$. Para cada valor de r (comprendido entre 0 y el menor de las m y n), todas las matrices de rango r forman una clase de equivalencia (formada por matrices equivalentes entre sí); a esta clase pertenece la matriz C_r, que se tomará como representante de la misma (representante canónico de equivalencia).

EJERCICIO

Sea $f: \mathbb{R}^4 \to \mathbb{R}^3$ la aplicación lineal que, respecto de ciertas bases:

$$(\bar{e}_1, \bar{e}_2, \bar{e}_3, \bar{e}_4) \text{ de } \mathbb{R}^4 \quad \text{y} \quad (\bar{u}_1, \bar{u}_2, \bar{u}_3) \text{ de } \mathbb{R}^3$$

tiene asociada la siguiente matriz A:

$$A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ -5 & -9 & -4 & 0 \\ -8 & -14 & -8 & -2 \end{bmatrix}$$

Se pide:

1. Una base de $\mathbb{R}^4$ y otra de $\mathbb{R}^3$ respecto de las cuales f tenga asociada la matriz canónica de equivalencia.
2. Bases del núcleo y de la imagen de f.

RESOLUCION

1. Realizando adecuadas operaciones elementales en las filas de la matriz $[A \mid I_3]$ (*) se obtiene fácilmente una matriz $[A' \mid Q^{-1}]$, donde A' es triangular (escalonada) superior:

$$[A' \mid Q^{-1}] = \left[\begin{array}{cccc|ccc} 1 & 2 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -4 & -5 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & -2 & -2 & 1 \end{array}\right] \; ; \; Q = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ -8 & 2 & 1 \end{bmatrix}$$

Realizando ahora operaciones elementales en las filas de la matriz $[A'/I_4]$ (**) se obtiene fácilmente una matriz $[C_r/P]$, donde C_r es la matriz canónica de equivalencia de A:

$$\left[\frac{A'}{I_4}\right] = \left[\begin{array}{cccc} 1 & 2 & 0 & -1 \\ 0 & 1 & -4 & -5 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \cong \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & -2 & -8 & -9 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \left[\frac{C_r}{P}\right]$$

Conocidas P y Q (que no son únicas; con otras operaciones elementales se obtendrían unas P y Q distintas), unas bases de $\mathbb{R}^4$ y $\mathbb{R}^3$ en las que f tiene por matriz a C_r son:

(*) Una matriz que resulta de escribir A y poner a su derecha la matriz unidad 3×3.
(**) Matriz que resulta de escribir A' y poner debajo de ella la matriz unidad 4×4.

Base de $\mathbb{R}^4$	Base de $\mathbb{R}^3$
$\bar{e}'_1 = \bar{e}_1$	$\bar{u}'_1 = \bar{u}_1 - 5\bar{u}_2 - 8\bar{u}_3$
$\bar{e}'_2 = 2\bar{e}_1 + \bar{e}_2$	$\bar{u}'_2 = \bar{u}_2 - 2\bar{u}_3$
$\bar{e}'_3 = -8\bar{e}_1 + 4\bar{e}_2 + \bar{e}_3$	$\bar{u}'_3 = \bar{u}_3$
$\bar{e}'_4 = -9\bar{e}_1 + 5\bar{e}_2 + \bar{e}_4$	

2. Como $\operatorname{rang} f = \operatorname{rang} A = 2$, resulta que la imagen de f tiene dimensión $r = 2$ y el núcleo de f tiene dimensión $n - r = 4 - 2 = 2$. Por ello, las bases pedidas son:

   - Una base de $\operatorname{Im}(f) = (\text{los } r \text{ primeros vectores de la nueva base de } \mathbb{R}^3) = (\bar{u}'_1, \bar{u}'_2)$
   - Una base de $\operatorname{Nuc}(f) = (\text{los } n - r \text{ últimos vectores de la nueva base de } \mathbb{R}^4) = (\bar{e}'_4, \bar{e}'_3)$.

MATRICES DE UN ENDOMORFISMO. MATRICES SEMEJANTES

Para acabar este asunto de las matrices asociadas a una aplicación lineal $f: V \to W$, vamos a analizar el caso en el que $V = W$, es decir, supondremos que f es un endomorfismo, de V en V. Ahora, en lugar de manejar dos bases, una en V y otra en W, únicamente habrá una, a la que se referirán todos los vectores: los $\bar{x} \in V$ y los $f(\bar{x}) \in V$.

[094]

Sea $f: V \to V$ un endomorfismo, en el espacio vectorial V de dimensión finita n sobre un cuerpo K, y sea A la matriz asociada a f respecto de cierta base de V. Si en V se cambia la base y llamando P a la matriz del respectivo cambio de coordenadas, entonces la matriz asociada a f en la nueva base es la $A' = P^{-1}AP$.

Dos matrices cuadradas A y A', del mismo tamaño $n \times n$, se dicen *matrices semejantes* si están asociadas a un mismo endomorfismo de K^n en K^{n} (respecto de bases adecuadas), lo que equivale a que exista una matriz regular P, de tamaño $n \times n$, tal que $A' = P^{-1}AP$.

DEMOSTRACION

Las anteriores afirmaciones se comprueban como se hizo en [090]. Concretamente, basta con suponer que ahora es $W = V$ y $Q = P$; con ello, lo allí demostrado es de aplicación a este caso.

EJEMPLO

Sea $f:\mathbb{R}^3 \to \mathbb{R}^3$ el endomorfismo que respecto de la base canónica, tiene asociada la matriz

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 3 \end{bmatrix}$$

Hallar la matriz asociada a f en la base $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$, donde

$$\bar{u}_1 = (1, 1, 0) \quad ; \quad \bar{u}_2 = (1, 0, 1) \quad ; \quad \bar{u}_3 = (0, 1, 1)$$

RESOLUCION

La matriz P de cambio de coordenadas y su inversa P^{-1} (que se obtiene fácilmente) son:

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \; ; \; P^{-1} = \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ -1/2 & 1/2 & 1/2 \end{bmatrix}$$

Por tanto, la matriz de f en la nueva base será la

$$A' = P^{-1}AP = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

DETERMINANTE DE UN ENDOMORFISMO

[095]

Para cualquiera que sea el endomorfismo $f:V \to V$, de un espacio vectorial V de dimensión finita, todas las matrices asociadas a f en las distintas bases de V tienen el mismo determinante, que se llama *determinante del endomorfismo* f y se representa por $\det f$. Este resultado puede enunciarse, de modo equivalente, diciendo que dos matrices cuadradas que sean semejantes tienen el mismo determinante.

Sean $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ y $B' = (\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n)$ dos bases de un mismo espacio vectorial V. Se dice que B y B' son dos *bases con la misma o con distinta orientación* si es positivo o negativo, respectivamente, el determinante del automorfismo f que transforma B en B', esto es aquel para el que $f(\bar{e}_i) = \bar{e}'_i$ para $i = 1, 2, ..., n$. Las bases de V se dividen en dos clases, que tienen orientaciones distintas.

DEMOSTRACION

1. Si A y A' son matrices de un mismo endomorfismo $f: V \to V$ o si A y A' son semejantes, sabemos que existe una matriz regular P tal que $A' = P^{-1}AP$. Tomando determinantes en esta igualdad, se obtiene:

$$\det A' = (\det P)^{-1}(\det A)(\det P) = \det A$$

2. Para el conjunto de todas las bases de V, la relación $\simeq$ definida por

$$(B \simeq B') \Leftrightarrow (B \text{ y } B' \text{ tienen la misma orientación})$$

en una equivalencia. En efecto, llamando (B, B') al endomorfismo que transforma la base B en la base B', se tiene:

- $B \simeq B \ \forall B$ ya que $(B, B) = i$ (identidad) y $\det i = 1 > 0$.
- $B \simeq B' \Rightarrow B' \simeq B$, ya que $(B', B) = (B, B')^{-1}$ y, como $\det(B, B') > 0$, será $\det(B', B) = [\det(B, B')]^{-1} > 0$.
- $B \simeq B'$ y $B' \simeq B'' \Rightarrow B \simeq B''$, ya que $(B, B'') = (B', B'') \circ (B, B')$, luego si $\det(B, B') > 0$ y $\det(B', B'') > 0$ será

$$\det(B, B'') = [\det(B', B'')][\det(B, B')] > 0$$

En esta relación de equivalencia hay dos clases. En efecto: sea B_1 una base cualquiera y llamemos B_2 a la base que resulte de cambiar el primer vector $\vec{e}_1$ de B_1 por $-\vec{e}_1$. Es evidente que B_1 y B_2 son de distinta clase, pues $\det(B_1, B_2) = \det(-I) = -1$. Para cualquier otra base B, como

$$\det(B, B_2) = [\det(B_1, B_2)][\det(B, B_1)] = -\det(B, B_1)$$

resulta que B tiene la misma orientación que una de las B_1 o B_2 y distinta orientación que la otra.

γ OPERACIONES CON APLICACIONES LINEALES

Las operaciones de las que nos vamos a ocupar aquí son las usuales entre funciones en general. Hablaremos de la suma de dos funciones, del producto de un escalar por una función y de la composición de dos funciones. Acontece que cuando se suman, cuando se multiplican por escalares y cuando se componen funciones que son lineales, se obtienen nuevas funciones lineales.

Este asunto se analizará con detalle; se van a estudiar las estructuras algebraicas que con las susodichas operaciones, se originan en los conjuntos de las aplicaciones lineales. En concreto, se analizarán la conjunta $\mathcal{L}(V, W)$, de los homomorfismos del espacio V en el W, el conjunto $\mathcal{L}(V)$, de los endomorfismos

de V, y el conjunto $GL(V)$, de los automorfismos de V. El primero resulta ser un espacio vectorial; el segundo, un anillo; el tercero, un grupo.

Todo lo anterior cobra, para nosotros, especial relevancia si V y W tienen dimensión finita. En estos casos, recurriendo a las matrices asociadas a las aplicaciones lineales, se establece un paralelismo entre las estructuraciones algebraicas de los homomorfismos y de las matrices. Mejor dicho, se definen las operaciones entre matrices de modo que exista isomorfismo entre cada estructura algebraica de homomorfismo y su correspondiente estructura de matrices.

6.5. ESPACIOS VECTORIALES DE HOMOMORFISMOS Y DE MATRICES

En primer lugar consideraremos el conjunto de las aplicaciones lineales entre dos espacios V y W distintos, en general, dedicando particular atención al caso de dimensión finita, en el que intervienen las matrices.

ESPACIO VECTORIAL $\mathcal{L}(V, W)$

[096]

Se denota por $\mathcal{L}(V, W)$(*) al conjunto de las aplicaciones lineales del espacio V en el W, ambos sobre un mismo cuerpo K. Se consideran la suma $f+g$ y el producto λf usuales, para cualesquiera f y g de $\mathcal{L}(V, W)$ y todo $\lambda \in K$:

$$(f+g)(\bar{u}) = f(\bar{u}) + g(\bar{u}) \quad , \quad (\lambda f)(\bar{u}) = \lambda f(\bar{u}) \quad (\forall\, \bar{u} \in V)$$

Las aplicaciones $f+g$ y λf son lineales. Estas operaciones confieren a $\mathcal{L}(V, W)$ estructura de espacio vectorial, sobre el cuerpo K. Esto es, para cualesquiera $f, g, h \in \mathcal{L}(V, W)$ y $\lambda, \mu \in K$, se verifica:

$$\begin{array}{c|c} (f+g)+h = f+(g+h) & \lambda(f+g) = \lambda f + \lambda g \\ f+o = o+f = f & (\lambda+\mu)f = \lambda f + \mu f \\ f+(-f) = o & (\lambda\mu)f = \lambda(\mu f) \\ f+g = g+f & 1f = f \end{array}$$

(o es la aplicación nula. La aplicación $-f$, o sea $(-f)(\bar{u}) = -f(\bar{u})$, $\forall\, \bar{u} \in V$, es lineal.)

(*) También se suele poner $\mathrm{Hom}(V, W)$.

DEMOSTRACION

- Las funciones $f+g$ y λf son lineales ya que, para cualesquiera $\vec{u}_1, \vec{u}_2 \in V$ y $\alpha_1, \alpha_2 \in K$, se verifica:

$$(f+g)(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2) = f(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2) + g(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2) = [\alpha_1 f(\vec{u}_1) + \alpha_2 f(\vec{u}_2)] + [\alpha_1 g(\vec{u}_1) + \alpha_2 g(\vec{u}_2)] = \alpha_1[f(\vec{u}_1) + g(\vec{u}_1)] + \alpha_2[f(\vec{u}_2) + g(\vec{u}_2)] = \alpha_1(f+g)(\vec{u}_1) + \alpha_2(f+g)(\vec{u}_2)$$

$$(\lambda f)(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2) = \lambda f(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2) = \lambda[\alpha_1 f(\vec{u}_1) + \alpha_2 f(\vec{u}_2)] = \alpha_1 \lambda f(\vec{u}_1) + \alpha_2 \lambda f(\vec{u}_2) = \alpha_1(\lambda f)(\vec{u}_1) + \alpha_2(\lambda f)(\vec{u}_2)$$

- Para comprobar que, con estas operaciones, $\mathcal{L}(V, W)$ es un subespacio del espacio $\mathcal{F}(V, W)$, de las aplicaciones cualesquiera de V en W. Esta comprobación se reduce a ver que, si dos aplicaciones f_1 y f_2 de V en W, son lineales, también es lineal la $\alpha_1 f_1 + \alpha_2 f_2$, para cualesquiera $\alpha_1, \alpha_2 \in K$; esto es justamente lo que acabamos de comprobar.

ESPACIO VECTORIAL $\mathcal{M}_{m\times n}$

[097]

Se denota por $\mathcal{M}_{m\times n}$ (*) al conjunto de las matrices de tamaño $m \times n$, con elementos de un cierto cuerpo K; cada una de ellas, A, se corresponde biyectivamente con su aplicación lineal asociada f_A (de K^n en K^m; respecto de las bases canónicas). La suma $A+B$ y el producto λA, para cualesquiera $A, B \in \mathcal{M}_{m\times n}$ y $\lambda \in K$, son las matrices de $\mathcal{M}_{m\times n}$ definidas de uno cualquiera de los dos modos siguientes(**) (equivalentes entre sí):

1. Si $A = [a_{ij}]$ y $B = [b_{ij}]$, entonces $A + B = [a_{ij} + b_{ij}]$ y $\lambda A = [\lambda a_{ij}]$.
2. $A+B$ y λA son las matrices asociadas a las aplicaciones lineales $f_A + f_B$ y λf_A, donde f_A y f_B son las aplicaciones lineales asociadas a A y B.

Estas operaciones confieren a $\mathcal{M}_{m\times n}$ estructura de espacio vectorial isomorfo al $\mathcal{L}(K^n, K^m)$, de manera que; para cualesquiera $A, B, C \in \mathcal{M}_{m\times n}$ y $\lambda, \mu \in K$, se verifica:

$(A+B)+C = A+(B+C)$	$\lambda(A+B) = \lambda A + \lambda B$
$A+O = O+A = A$	$(\lambda+\mu)A = \lambda A + \mu A$
$A+(-A) = O$	$(\lambda\mu)A = \lambda(\mu A)$
$A+B = B+A$	$1A = A$

(O es la matriz nula. La matriz $-A = -[a_{ij}] = [-a_{ij}]$ es la opuesta de A.)

(*) Para explicitar el cuerpo K de escalares, en lugar de $\mathcal{M}_{m\times n}$ se pondría $\mathcal{M}_{m\times n}(K)$.
(**) El primero de ellos ya fue adelantado en [022].

DEMOSTRACION

- Las aplicaciones lineales f_A y f_B son las que tienen por ecuaciones a

$$f_A : (x_j) \longmapsto (y_i) = (\Sigma\, a_{ij} x_j) \quad \text{y} \quad f_B : (x_j) \longmapsto (y_i) = (\Sigma\, b_{ij} x_j)$$

Por ello, las ecuaciones de $f_A + f_B$ y λf_A son:

$$f_A + f_B : (x_j) \longmapsto (\Sigma\, (a_{ij} + b_{ij}) x_j) \quad \text{y} \quad \lambda f_A : (x_j) \longmapsto (\Sigma\, (\lambda a_{ij}) x_j)$$

Resulta, pues, que las matrices asociadas a $f_A + f_B$ y λf_A son:

$$[\text{matriz de } f_A + f_B] = [a_{ij} + b_{ij}] \quad \text{y} \quad [\text{matriz de } \lambda f_A] = [\lambda a_{ij}]$$

Esto confirma que las dos definiciones de suma y de producto por un escalar coinciden.

- El segundo de los dos modos de definir las anteriores operaciones entre matrices permite asegurar que la biyección entre $\mathcal{M}_{m\times n}$ y $\mathcal{L}(K^n, K^m)$ es un isomorfismo y, como $\mathcal{L}(K^n, K^m)$ es un espacio vectorial, de ello se desprende que también lo es $\mathcal{M}_{m\times n}$.

DIMENSION DE $\mathcal{L}(K^n, K^m)$ Y DE $\mathcal{M}_{m\times n}$

[098]

Los espacios vectoriales $\mathcal{L}(K^n, K^m)$ y $\mathcal{M}_{m\times n}$ tienen, ambos, dimensión $m \cdot n$. Una base del segundo de ellos está formada por las $m \cdot n$ matrices E_{ij} (para $i = 1, 2, ..., m$ y $j = 1, 2, ..., n$) siguientes:

E_{ij} = matriz de tamaño $m \times n$ que tiene todos sus elementos nulos excepto el que ocupa el lugar ij, que es igual a la unidad.

(A esta base la llamaremos base usual de $\mathcal{M}_{m\times n}$.)

DEMOSTRACION

El espacio $\mathcal{M}_{m\times n}$ tiene dimensión $m \cdot n$ y una base suya es formada por las E_{ij}, según ya se comprobó en el ejemplo 2 de [061]. El espacio $\mathcal{L}(K^n, K^m)$ también tiene dimensión $m \cdot n$, ya que es isomorfo al $\mathcal{M}_{m\times n}$.

EJERCICIO

Comprobar que las siguientes aplicaciones lineales f, g y h, de $\mathbb{R}^3$ en $\mathbb{R}^2$, son (como elementos del espacio vectorial $\mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$) linealmente dependientes

$$f(x, y, z) = (x - y + 2z,\ 3x + y)$$
$$g(x, y, z) = (3x + y,\ -x + 5y + 4z)$$
$$h(x, y, z) = (2x + z,\ x + 3y + 2z)$$

RESOLUCION

Las aplicaciones f, g y h serán linealmente dependientes si y sólo si lo son sus matrices asociadas:

$$M_f = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 0 \end{bmatrix} \; ; \; M_g = \begin{bmatrix} 3 & 1 & 0 \\ -1 & 5 & 4 \end{bmatrix} \; ; \; M_h = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 3 & 2 \end{bmatrix}$$

Estas tres matrices son linealmente dependientes si y sólo si lo son sus vectores de coordenadas, respecto de la base usual, que son:

$$\bar{v}_f = (1, -1, 2, 3, 1, 0)$$
$$\bar{v}_g = (3, 1, 0, -1, 5, 4)$$
$$\bar{v}_g = (2, 0, 1, 1, 3, 2)$$

Estos vectores son linealmente dependientes, pues $\bar{v}_f + \bar{v}_g = 2\bar{v}_h$, luego también lo son f, g y h.

6.6. ANILLO DE ENDOMORFISMOS. MATRICES INVERTIBLES

COMPOSICION DE HOMOMOI:FISMOS Y PRODUCTO DE MATRICES

Tanto la composición de aplicaciones lineales (véase [081], 3) como el producto de matrices (véase [023]) nos son ya conocidos. Recogemos aquí estas cuestiones para relacionarlas entre sí, justificando así la definición de producto de matrices que, tal y como está en este momento, parece un tanto caprichosa.

[099]

Dados tres espacios vectoriales V, W y U, los tres sobre el mismo cuerpo K, toda aplicación lineal $g: V \to W$ se puede componer con cualquier aplicación lineal $f: W \to U$, siendo el resultado $f \circ g: V \to U$ una nueva aplicación lineal.

Siempre que las siguientes composiciones tengan sentido, se verifica que (f, g y h homomorfismos; λ escalar):

$$(f \circ g) \circ h = f \circ (g \circ h) \quad (1)$$
$$(f + g) \circ h = f \circ h + g \circ h \quad (2)$$
$$h \circ (f + g) = h \circ f + h \circ g \quad (3)$$
$$\lambda(f \circ g) = (\lambda f) \circ g = f \circ (\lambda g) \quad (4)$$

DEMOSTRACION

En [081] 3) ya se comprobó que la composición de dos aplicaciones lineales es otra aplicación lineal. Sólo hay, así, que probar las cuatro igualdades del enunciado. Estas son ciertas ya que, para cualquiera que sea el vector $\bar{x}$ al que se le aplique unas y otras composiciones de funciones, se verifica que (se omite la (3) por analogía con la (2)):

$$\left.\begin{aligned}[(f\circ g)\circ h](x) &= (f\circ g)(h(x)) = f(g(h(x)))\\ [f\circ(g\circ h)](x) &= f[(g\circ h)(x)] = f(g(h(x)))\end{aligned}\right\} \Rightarrow (1)$$

$$\left.\begin{aligned}[(f+g)\circ h](x) &= (f+g)(h(x)) = f(h(x)) + g(h(x))\\ [f\circ h + g\circ h](x) &= (f\circ h)(x) + (g\circ h)(x) = f(h(x)) + g(h(x))\end{aligned}\right\} \Rightarrow (2)$$

$$\left.\begin{aligned}[\lambda(f\circ g)](x) &= \lambda(f\circ g)(x) = \lambda f(g(x))\\ [(\lambda f)\circ g](x) &= (\lambda f)(g(x)) = \lambda f(g(x))\end{aligned}\right\} \Rightarrow (4)$$

[100]

Dadas las matrices $A \in \mathcal{M}_{m\times p}$ y $B \in \mathcal{M}_{p\times n}$ y haciendo corresponder cada una de ellas con su aplicación lineal asociada ($f_A: K^p \to K^m$ y $f_B: K^n \to K^p$, respecto de las bases canónicas), se llama producto $A \cdot B$ o AB a la matriz de $\mathcal{M}_{m\times n}$ definida de uno cualquiera de los dos modos, equivalentes entre sí:

1. Si $A = [a_{ih}]$ y $B = [b_{hj}]$, entonces $AB = [c_{ij}]$ con $c_{ij} = \sum_{h=1}^{p} a_{ih}b_{hj}$ (*).
2. AB es la matriz asociada a la aplicación lineal $f_A \circ f_B$.

Siempre que los siguientes productos tengan sentido, se verifica que, si A, B y C son matrices y λ es escalar:

$$(AB)C = A(BC)$$
$$(A+B)C = AC + BC$$
$$C(A+B) = CA + CB$$
$$\lambda(AB) = (\lambda A)B = A(\lambda B)$$

(*) $i = 1, 2, \ldots, m$ y $j = 1, 2, \ldots, n$. Esta definición ya ha sido considerada en [023]; allí se comenta y desarrolla extensamente, se analizan los casos más interesantes y se hacen observaciones al respecto.

DEMOSTRACION

• Las aplicaciones lineales f_A y f_B son las que tienen por ecuaciones a:

$$f_B: (x_j) \mapsto (y_h) = \left(\sum_j b_{hj}x_j\right) \quad \text{para } j = 1, 2, \ldots, n \quad h = 1, 2, \ldots, p$$
$$f_A: (y_h) \mapsto (z_i) = \left(\sum_h a_{ih}y_h\right) \quad i = 1, 2, \ldots, m$$

Por ello, las ecuaciones de $f_A \circ f_B$ son:

$$f_A \circ f_B : (x_j) \mapsto (z_i) = \left(\sum_h a_{ih}\left(\sum_j b_{hj}x_j\right)\right) = \left(\sum_j \left(\sum_h a_{ih}b_{hj}\right)x_j\right)$$

Resulta, pues, que la matriz asociada a $f_A \circ f_B$ es:

$$[\text{Matriz de } f_A \circ f_B] = [c_{ij}], \quad \text{donde} \quad c_{ij} = \sum_h a_{ih}b_{hj}$$

lo cual prueba que las dos definiciones dadas de producto son, en efecto, coincidentes.

- Las cuatro igualdades que figuran al final de este enunciado son consecuencia obvia de las correspondientes igualdades, entre aplicaciones lineales, que acabamos de considerar en el anterior enunciado [099]. Nótese a este respecto que las biyecciones

$$\mathcal{L}(K^r, K^s) \mapsto \mathcal{M}_{s\times r}, \quad A \mapsto f_A$$

se convierten en isomorfismos al considerar la composición de aplicaciones en $\mathcal{L}(K^r, K^s)$ y el producto de matrices en $\mathcal{M}_{s\times r}$. Recuérdese que, por otra parte, las propiedades del producto de matrices fueron ya estudiadas anteriormente (véase [024]).

EJERCICIO

Sean $f : \mathbb{R}^3 \to \mathbb{R}^2$ y $g : \mathbb{R}^2 \to \mathbb{R}^3$ las aplicaciones lineales que tienen por ecuaciones a:

$$f : (x, y, z) \mapsto f(x, y, z) = (x + 2y + 3z, -x + y + 5z)$$
$$g : (u, v) \mapsto g(u, v) = (u + v, 2u - v, 3u - 4v)$$

Hallar las ecuaciones matriciales (en bases canónicas) de $g \circ f$ y de $f \circ g$ (nótese que, en este caso, ambas composiciones tienen sentido).

RESOLUCION

Las matrices de f y g son las

$$M_f = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1 & 5 \end{bmatrix} \quad \text{y} \quad M_g = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 3 & -4 \end{bmatrix}$$

Por tanto, las ecuaciones pedidas son

$$g \circ f : (x, y, z) \mapsto M_g \cdot M_f \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 3 & 8 \\ 3 & 3 & 1 \\ 7 & 2 & -11 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$f \circ g : (u, v) \mapsto M_f \cdot M_g \cdot \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 14 & -13 \\ 16 & -22 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

ANILLOS DE ENDOMORFISMOS Y DE MATRICES

Un caso particular del anterior, que tiene especial relevancia, es aquel en el que las aplicaciones lineales que se componen lo son de un espacio V en sí mismo (endomorfismos en V) o, si hablamos del producto de matrices, aquel en el que se multiplican matrices cuadradas, del mismo tamaño. En estos supuestos, la composición de endomorfismos o el producto de matrices cuadradas son operaciones internas para el conjunto, de endomorfismos o de matrices cuadradas, que ahora se considera. Con la suma y con esta nueva operación de la que estamos hablando, se obtienen estructuras algebraicas de anillo, de las que pasamos a ocuparnos.

[101]

Se denotará por $\mathcal{L}(V)$(*) al conjunto de los endomorfismos de V; la suma y la composición de aplicaciones son operaciones internas para $\mathcal{L}(V)$. Ellas le confieren estructura de anillo(**) unitario, no conmutativo y con divisores de cero; esto es, $\mathcal{L}(V)$ es grupo abeliano para la suma y, además, se verifica (para cualesquiera $f, g, h \in \mathcal{L}(V)$):

$(f\circ g)\circ h = f\circ(g\circ h)$	$f\circ i = i\circ f = f$ (i = aplicación idéntica)
$f\circ(g+h) = f\circ g + f\circ h$	$f\circ g \neq g\circ f$ (en general)
$(g+h)\circ f = g\circ f + h\circ f$	$f\circ g = o \not\Rightarrow [f = o \text{ ó } g = o]$ (en general)

(*) Con la notación de [096], $\mathcal{L}(V)$ no es otra cosa que $\mathcal{L}(V, V)$; también se le suele denotar poniendo End (V).

(**) Véase el Apéndice A.9 al final del libro.

DEMOSTRACION

El que la suma y la composición de aplicaciones son operaciones internas para $\mathcal{L}(V)$ son consecuencias obvias de lo dicho en [096] y [099]. Respecto de las igualdades anteriores, las tres primeras ya se consideraron en [099] (para un caso más general), la cuarta es evidente. Para verificar las dos últimas relaciones, bastará con poner sendos ejemplos:

1. Sea V el espacio vectorial de los polinomios $p(x)$ de coeficientes reales; sean $f, g \in \mathcal{L}(V)$ los endomorfismos:

$$f : p(x) \mapsto f(p(x)) = p(-x)$$
$$g : p(x) \mapsto g(p(x)) = p'(x)$$

Las composiciones $f\circ g$ y $g\circ f$ son distintas ya que:

$$f[g(a_0 + a_1x + a_2x^2 + a_3x^3 z + \cdots + a_nx^n)] = a_1 - 2a_2x + 3a_3x^2 - \pm na_nx^{n-1}$$
$$g[f(a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots a_nx^n)] = -a_1 + 2a_2x - 3a_3x^2 + \cdots \mp na_nx^{n-1}$$

son dos polinomios distintos.

2. Sea V el espacio vectorial de las sucesiones indefinidas de números reales(*) y considérense las siguientes aplicaciones $f: V \to V$ y $g: V \to V$, que evidentemente son lineales:

$$f(x_1, x_2, x_3, x_4, ..., x_{2n-1}, x_{2n}, ...) = (x_1, 0, x_3, 0, ..., x_{2n+1}, 0, ...)$$

$$g(x_1, x_2, x_3, x_4, ..., x_{2n-1}, x_{2n}, ...) = (0, x_2, 0, x_4, ..., 0, x_{2n}, ...)$$

Ninguna de ellas es el endomorfismo nulo y, sin embargo, las dos composiciones $g \circ f$ y $f \circ g$ son nulas.

[102]

Se denota por $\mathcal{M}_n$ (1) al conjunto de las matrices cuadradas de tamaño n, con elementos de un cierto cuerpo K; la suma y el producto de matrices son operaciones internas para $\mathcal{M}_n$. Ellas le confieren estructura de anillo(2) unitario, no conmutativo y con divisores de cero; esto es, $\mathcal{M}_n$ es un grupo abeliano para la suma y, además, se verifica (para cualesquiera $A, B, C \in \mathcal{M}_n$):

$(AB)C = A(BC)$	$AI = IA = A$ (I = matriz unidad)
$A(B+C) = AB + AC$	$AB \neq BA$ (en general)
$(B+C)A = BA + CA$	$AB = O \not\Rightarrow [A = O \text{ o } B = O]$ (en general)

Este anillo $\mathcal{M}_n$ es isomorfo al $\mathcal{L}(K^n)$, de los endomorfismos de K^n.

(1) $\mathcal{M}_n$ no es otra cosa que $\mathcal{M}_{n\times n}$ (véase [097]). Si hubiera necesidad de señalar cuál era el cuerpo K de escalares, en lugar de $\mathcal{M}_n$, pondría $\mathcal{M}_n(K)$.

(2) Véase el Apéndice A.9 al final del libro.

Recibe el nombre de *álgebra* la siguiente estructura algebraica: un conjunto A dotado de dos operaciones internas, que llamaremos suma (+) y producto($\cdot$), y de un producto por escalares (elementos de un cierto cuerpo K) tales que: 1) A con la suma y el producto es un anillo; 2) A con la suma y el producto por escalares es un espacio vectorial, y 3) para cualesquiera $\lambda \in K$ y $a, a' \in A$, se verifica que $\lambda(a \cdot a') = (\lambda a) \cdot a' = a \cdot (\lambda a')$. Los conjuntos $\mathcal{L}(V)$ y $\mathcal{M}_n$, de los endomorfismos en un espacio vectorial V y de las matrices cuadradas de tamaño n, son álgebras respecto de las operaciones usuales. También es un álgebra el conjunto de los polinomios con coeficientes en un cierto cuerpo.

DEMOSTRACION

La biyección, entre $\mathcal{M}_n$ y $\mathcal{L}(K^n)$, que a cada matriz le hace corresponder su endomorfismo asociado, respecto de la base canónica, es lineal respecto de la suma y del producto-composición, ya que de acuerdo con las definiciones de estas operaciones (véase [097] y [100]), para cualesquiera $A, B \in \mathcal{M}_n$, es:

$$[\text{Endomorfismo de } A + B] = [\text{End. de } A] + [\text{End. de } B]$$

$$[\text{Endomorfismo de } AB] = [\text{End. de } A] \circ [\text{End. de } B]$$

(*) En este espacio, la suma y el producto por escalar se definen, como para sucesiones limitadas, mediante:

$$(x_1, x_2, ..., x_n, ...) + (y_1, y_2, ..., y_n, ...) = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n, ...)$$

$$\lambda(x_1, x_2, ..., x_n, ...) = (\lambda x_1, \lambda x_2, ..., \lambda x_n, ...)$$

Por tanto, dicha biyección es un isomorfismo y, como $\mathcal{L}(K^n)$ es anillo respecto de la suma y la composición, también los $\mathcal{M}_n$ respecto de la suma y el producto. Como el primer anillo es unitario, también lo es el segundo. Las dos últimas relaciones se verifican fácilmente con algunos ejemplos, como los siguientes:

1. Para matrices cuadradas A y B, en general es $AB \neq BA$, como en el caso:

$$AB = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} -7 & -3 \\ 8 & 4 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & -6 \\ 1 & -5 \end{bmatrix}$$

2. Para matrices cuadradas A y B, el producto AB puede ser nulo sin que lo sean ni A ni B, como en el caso:

$$\begin{bmatrix} 1 & -2 \\ -3 & 6 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

AUTOMORFISMOS Y MATRICES REGULARES

Recordemos que en un anillo, como los $\mathcal{L}(V)$ y $\mathcal{M}_n$ recién analizados, los elementos que tienen inverso desempeñan un papel destacado; el conjunto que forman los elementos con inverso es un grupo respecto del producto, que se llama grupo multiplicativo del anillo (véase [398], en el Apéndice 9). Vamos ahora a ocuparnos de dichos elementos que son: en $\mathcal{L}(V)$, los automorfismos de V o aplicaciones lineales biyectivas de V en V; y en $\mathcal{M}_n$, las matrices cuadradas regulares o invertibles.

[103]

Se denota por $GL(V)$ al conjunto de los automorfismos de un espacio vectorial V, es decir, de los endomorfismos $f: V \to V$ que son biyectivos. Este conjunto $GL(V)$ es un grupo respecto de la composición de aplicaciones, que recibe el nombre de *grupo lineal* de V.

Se denota por $GL(n)$ al conjunto de las matrices cuadradas regulares(*) (de escalares de cierto cuerpo K). Este conjunto $GL(n)$ es un grupo respecto del producto de matrices, que recibe el nombre de *grupo lineal* de tamaño n.

Los grupos $GL(K^n)$ y $GL(n)$ son isomorfos.

(*) Recuérdese que una matriz cuadrada A de tamaño $n \times n$ es regular si se cumple una cualquiera de las siguientes condiciones equivalentes entre sí: A tiene inversa, A es simplificable para el producto, A no es divisor de cero, $\det A \neq 0$ (véase [027], [030] y [040]).

DEMOSTRACION

Los conjuntos $GL(V)$ y $GL(n)$ son grupos por tratarse de los conjuntos que forman los elementos inversibles de los anillos $\mathcal{L}(V)$ y $\mathcal{M}_n$ respectivamente (véase [398], en el Apéndice 9). Los grupos $GL(K^n)$ y $GL(n)$ son isomorfos ya que los anillos $\mathcal{L}(K^n)$ y $\mathcal{M}_n$ son isomorfos.

EJEMPLO

Seas V el espacio vectorial de los polinomios $p(x)$ de coeficientes reales y considérese la aplicación

$$f: V \to V \quad , \quad p(x) \mapsto f(p(x)) = p(x) + xp'(x)$$

Nótese que la expresión de $f(p(x))$ es:

$$f(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_0 + 2a_1x + 3a_2x^2 + \cdots + (n+1)a_nx^n$$

Esta aplicación f es evidentemente lineal y biyectiva; se trata, pues, de un automorfismo de V, esto es, $f \in GL(V)$. El automorfismo recíproco (o inverso) del f es el f^{-1} definido por

$$f^{-1}(b_0 + b_1x + b_2x^2 + \cdots + b_nx^n) = b_0 + \frac{b_1}{2}x + \frac{b_2}{3}x^2 + \cdots + \frac{b_n}{n+1}x^n$$

Nótese que, para cualquier polinomio $q(x) \in V$, se puede poner

$$f^{-1}(q(x)) = \frac{1}{x}\int_0^x q(t)\,dt \quad , \quad f^{-1}(q(0)) = q(0)$$

δ SISTEMAS DE ECUACIONES LINEALES

Este tema de los sistemas de ecuaciones lineales lo venimos arrastrando desde la primera página. No es casualidad. Hemos querido empezar el estudio del Algebra Lineal por donde es más fácil y procurándonos, lo antes posible, herramientas y útiles que nos facilitarán el estudio de esta disciplina. Por ello, en primer lugar, antes que ninguna otra cosa. nos las vimos con los sistemas de ecuaciones lineales. Pero, obviamente, no pudimos entonces abordar este asunto con la extensión debida, cosa ésta que hemos venido haciendo a medida que nos ha sido posible. Con lo que aquí vamos a considerar, daremos por concluido el estudio de los sistemas de ecuaciones lineales.

Todo cuanto decimos a continuación será en torno del teorema de Rouché, sobre el que ya dijimos algo en [049]; no obstante, lo que allí se obtuvo no pudo ser suficiente, por lo que nos vemos obligados a completarlo ahora. Nuestro objetivo es comprobar que las soluciones de un sistema de ecuaciones lineales (compatible) forman una «variedad afín». Por ello, empezaremos defi-

niendo estas variedades, comprobaremos después que, en una aplicación lineal, la imagen recíproca de un vector es una variedad lineal y acabaremos demostrando el Teorema de Rouché.

6.7. VARIEDADES AFINES

Las rectas y los planos, de $\mathbb{R}^3$, son las variedades afines que, a buen seguro, primero hemos conocido. El plano que pasa por $\bar{a} \in \mathbb{R}^3$ y tiene como vectores de dirección a los $\bar{u}_1$ y $\bar{u}_2$ (linealmente independientes), no es otra cosa que el conjunto de todos los $\bar{x} \in \mathbb{R}^3$ que se pueden poner en la forma:

$$\bar{x} = \bar{a} + \lambda_1 \bar{u}_1 + \lambda_2 \bar{u}_1^2 \quad \text{(al variar } \lambda_1, \lambda_2 \in \mathbb{R})$$

Obsérvese que si llamamos U al subespacio de $\mathbb{R}^3$ que engendran $\bar{u}_1$ y $\bar{u}_2$, el referido plano es el conjunto

$$\{\bar{x} = \bar{a} + \bar{u} \,/\, \bar{u} \in U\} \quad \text{(plano que pasa por } \bar{a} \text{ y tiene a } U \text{ por dirección)}$$

Este conjunto, que se llama variedad afín, se denota por $\bar{a} + U$. Pues bien, en general, se define:

[104]

En un espacio vectorial V, se llama *variedad afín* que pase por $\bar{a} \in V$ y que tiene por dirección al subespacio $U \subset V$, al siguiente conjunto de vectores, que se denota por $\bar{a} + U$:

$$\bar{a} + U = \{\bar{a} + \bar{u} \,/\, \bar{u} \in U\}$$

Dos variedades afines $\bar{a}_1 + U_1$ y $\bar{a}_2 + U_2$ son iguales si y sólo si se verifican las dos condiciones siguientes $\bar{a}_1 - \bar{a}_2 \in U_1 = U_2$.

Si U es un subespacio del espacio vectorial V, la relación entre vectores $\bar{v}$ y $\bar{v}'$ de V dada por $\bar{v} - \bar{v}' \in U$ es una equivalencia; el correspondiente conjunto cociente, que se denota por V/U está formado por las clases $[\bar{v}]$, que son las variedades afines $[\bar{v}] = \bar{v} + U$. Las operaciones de V son compatibles con esta equivalencia, lo que permite definir en V/U las operaciones suma y producto por escalar mediante $[\bar{u}] + [\bar{v}] = [\bar{u} + \bar{v}]$ y $\lambda[\bar{v}] = [\lambda \bar{v}]$. Con estas operaciones, V/U es un espacio vectorial que se llama *espacio vectorial cociente*.

DEMOSTRACION

Antes de entrar en materia, conviene notar que, de acuerdo con la definición, para cualquiera que sea el subespacio vectorial U, se verifica que $\bar{o} + U = U$ y si $\bar{a} \in U$, también es $\bar{a} + U = U$.

1. Supongamos en primer lugar que $\bar{a}_1 - \bar{a}_2 \in U_1 = U_2$. Entonces, será:

$$\begin{aligned} \bar{a}_1 + U_1 = \bar{a}_1 + U_2 &= \{\bar{a}_1 + \bar{u}_2 \,/\, \bar{u}_2 \in U_2\} = \{\bar{a}_2 + (\bar{a}_1 - \bar{a}_2) + \bar{u}_2 \,/\, \bar{u}_2 \in U_2\} = \\ &= \{\bar{a}_2 + [(\bar{a}_1 - \bar{a}_2) + \bar{u}_2] \,/\, (\bar{a}_1 - \bar{a}_2) + \bar{u}_2 \in U_2\} = \\ &= \{\bar{a}_2 + \bar{u}_2' \,/\, \bar{u}_2' \,/\, \bar{u}_2' \in U_2\} = \bar{a}_2 + U_2 \end{aligned}$$

2. Supongamos ahora que $\bar{a}_1 + U_1 = \bar{a}_2 + U_2$. Es evidente que esta igualdad equivale a la $(\bar{a}_1 - \bar{a}_0) + U_1 = (\bar{a}_2 - \bar{a}_0) + U_2$, donde $\bar{a}_0 \in V$ puede ser cualquiera; tomando $\bar{a}_0 = \bar{a}_1$, se obtiene que $U_1 = (\bar{a}_2 - \bar{a}_1) + U_2$. Como $\bar{o} \in U_1$, ha de existir un $\bar{u} \in U_2$ tal que $\bar{o} = (\bar{a}_2 - \bar{a}_1) + \bar{u}$, o sea, $\bar{a}_1 - \bar{a}_2 = \bar{u}$ y entonces resulta que $\bar{a}_1 - \bar{a}_2 \in U_2$.

Por otra parte, como ha resultado que $\bar{a}_2 - \bar{a}_1 \in U_2$, podemos asegurar que $(\bar{a}_2 - \bar{a}_1) + U_2 = U_2$ (ya que es obvio que, si $\bar{a} \in U$, entonces $\bar{a} + U = U$). Dado que anteriormente hemos obtenido que $U_1 = (\bar{a}_2 - \bar{a}_1) + U_2$, de esta igualdad y de la anterior se concluye que es $U_1 = U_2$, con lo que concluye la demostración.

IMAGEN RECIPROCA DE UN VECTOR

[105]

Sea $f: V \to W$ una aplicación lineal, entre los espacios vectoriales V y W, y considérese un vector fijo $\bar{b} \in W$. Se verifica que:

1. Si $\bar{b} \notin \mathrm{Im}(f)$, entonces $f^{-1}(\bar{b}) = \phi$.
2. Si $\bar{b} \in \mathrm{Im}(f)$, es decir, si existe algún $\bar{x}_0 \in V$ tal que $f(\bar{x}_0) = \bar{b}$, entonces $f^{-1}(\bar{b})$ es la siguiente variedad afín:

$$f^{-1}(\bar{b}) = \bar{x}_0 + \mathrm{Nuc}(f)$$

DEMOSTRACION

El apartado 1 es evidente. El apartado 2 es consecuencia de la siguiente cadena de equivalencias:

$$\bar{x} \in f^{-1}(\bar{b}) \Leftrightarrow \bar{b} = f(\bar{x}) \Leftrightarrow f(\bar{x}_0) = f(\bar{x}) \Leftrightarrow f(\bar{x} - \bar{x}_0) = \bar{o} \Leftrightarrow$$
$$\Leftrightarrow \bar{x} - \bar{x}_0 \in \mathrm{Nuc}(f) \Leftrightarrow \bar{x} \in \bar{x}_0 + \mathrm{Nuc}(f)$$

Sea $f: V \to W$ una aplicación lineal entre espacios vectoriales. El espacio vectorial cociente $W/\mathrm{Nuc}(f)$ está formado por las clases $[\bar{v}] = \bar{v} + \mathrm{Nuc}(f)$; esta clase $[\bar{v}]$ es el conjunto de todos los vectores de V que tienen la misma imagen que el vector $\bar{v} \in V$. Este espacio cociente es isomorfo al espacio imagen $\mathrm{Im}(f)$; la aplicación $[\bar{v}] \mapsto f(\bar{v})$ es un isomorfismo entre ellos (*teorema de isomorfismo*).

EJEMPLOS

1. Sea V el espacio vectorial de las funciones de $\mathbb{R}$ en $\mathbb{R}$ que son derivables y tienen derivada continua; sea W el espacio vectorial de las funciones de $\mathbb{R}$ en $\mathbb{R}$. La aplicación derivada

$$D: V \to W \quad , \quad f \mapsto D[f] = f'$$

es lineal. Dada una función $\varphi \in W$, respecto de su imagen recíproca por D, puede ocurrir:

- $D^{-1}[\varphi] = \phi$, si φ no es una función continua.

- $D^{-1}[\varphi] = \{f_0 + c \,/\, c \in \mathbb{R}\} = f_0 + \mathbb{R}$, donde f_0 es una primitiva de φ, en el caso de que φ sea una función continua.

Nótese que, en este segundo supuesto, la imagen recíproca $D^{-1}[\varphi]$ es el conjunto de todas las primitivas de φ, que es lo que se llama integral indefinida de φ.

2. Sea $f: \mathbb{R}^3 \to \mathbb{R}^2$ la aplicación lineal de ecuación

$$f(x, y, z) = (2x - y, -x + z)$$

La imagen recíproca del vector $(1, 3) \in \mathbb{R}^2$ está formada por todos los vectores $(x, y, z) \in \mathbb{R}^3$ tales que

$$\left.\begin{aligned} 2x - y &= 1 \\ -x + z &= 3 \end{aligned}\right\}$$

o sea

$$\begin{cases} x = \text{cualquiera} \\ y = -1 + 2x \\ z = 3 + x \end{cases}$$

Es decir,

$$\begin{aligned} f^{-1}(1, 3) &= \{(\alpha, -1 + 2\alpha, 3 + \alpha) \,/\, \alpha \in \mathbb{R}\} = \\ &= (0, -1, 3) + \{(\alpha, 2\alpha, \alpha) \,/\, \alpha \in \mathbb{R}\} = \\ &= (0, -1, 3) + \text{Nuc}\,(f) \end{aligned}$$

6.8. TEOREMA DE ROUCHE

Vamos a hablar aquí de un sistema de m ecuaciones lineales con n incógnitas $x_1, x_2, \ldots, x_n$; los coeficientes del sistema son elementos de un cierto cuerpo K, en el que se buscan las soluciones:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \quad \text{ó} \quad AX = B \text{ (matricialmente)}$$

donde $A = [a_{ij}]$, de tamaño $m \times n$, es la matriz de los coeficientes, $X = [x_i]$ la matriz columna de las n incógnitas y $B = [b_i]$ es la matriz columna $m \times 1$ de los términos independientes; $AB = [a_{ij} \,|\, b_i]$ se llama matriz ampliada.

[106]

Sean dadas $A \in \mathcal{M}_{m\times n}$ y $B \in \mathcal{M}_{m\times 1}$; sea $X \in \mathcal{M}_{n\times 1}$ desconocida (incógnita). Se verifica que (*teorema de Rouché*):

1. El sistema $AX = B$, de m ecuaciones lineales con n incógnitas, tiene alguna solución si y sólo si $\operatorname{rang} A\!B = \operatorname{rang} A$; el sistema tiene solución única si y sólo si $\operatorname{rang} A\!B = \operatorname{rang} A = n$.
2. Si el sistema es compatible, entonces sus soluciones forman la variedad afín:

$$\{\text{Soluciones } AX = B\} = X_0 + \{\text{Soluciones de } AX = O\}\ (*)$$

donde X_0 es una solución de $AX = B$ y donde el conjunto de las soluciones de la ecuación homogénea $AX = O$ es el núcleo de la aplicación lineal $X \rightarrowtail Y = AX$ (este núcleo tiene dimensión $n - r$).

(*) Al conjunto de las soluciones de un sistema se le suele llamar la «solución general» del mismo; si se considera una sola solución, se la llamará «solución particular». Con esta terminología, se dirá que la solución general de un sistema $AX = B$ es igual a una solución particular de él más la solución general del sistema homogéneo asociado, $AX = O$.

COMPROBACION

La parte 1 del teorema ya se comprobó anteriormente (véase [049]). Nos ocuparemos, pues, fundamentalmente de la parte 2, aunque algo diremos también de aquélla.

Al recurrir a la aplicación lineal $f : \mathbb{R}^n \to \mathbb{R}^m$ que está asociada (en las bases canónicas) a la matriz A, esto es, a la $X \rightarrowtail Y = AX$, queda de manifiesto que las soluciones de la ecuación $AX = B$ constituyen la imagen recíproca de B, esto es, $f^{-1}(B)$ (*) que, de acuerdo con lo dicho en [105], 2; es:

$$f^{-1}(B) = X_0 + \operatorname{Nuc}(f) = X_0 + \{X \,/\, AX = O\} =$$
$$= X_0 + \{\text{Soluciones de } AX = O\}$$

Por tanto, el conjunto de las soluciones es el que se decía en el enunciado.

Para que exista solución de $AX = B$ es necesario y suficiente que B pertenezca a la imagen de f, es decir, que la columna B sea combinación lineal de las columnas de A, o sea, que $A\!B$ tenga el mismo rango que A. Además, la solución será única si y sólo si, además de ser $\operatorname{rang} A\!B = \operatorname{rang} A$, se verifica que $\operatorname{Nuc}(f) = O$, o sea, $\operatorname{rang} A = \operatorname{rang} f = n - 0 = n$, con lo queda también comprobada la parte 1 del teorema.

(*) Para no complicar la notación, se pone B en lugar de $\bar{b} = (b_1, b_2, \ldots, b_n)$; esto es, se abusa de notación escribiendo matrices columna, como B, en lugar de vectores, como $\bar{b}$.

ε ESPACIO DUAL

Históricamente, en geometría plana la dualidad era entendida como una correspondencia que trasportaba las propiedades de los puntos convirtiéndolas en propiedades de las rectas y recíprocamente. Así, por ejemplo, «estar alineados» (los puntos) se convertía en «pasar por un punto» (las rectas). Con la dualidad, las propiedades lineales entre puntos daban lugar a propiedades lineales entre rectas.

Al tratar algebraicamente estas cuestiones, los puntos han dado lugar a vectores y las rectas, al recurrir a sus ecuaciones, han conducido a las funciones lineales. La dualidad pasa, así, a ser una correspondencia entre vectores y «formas lineales» que permite dotar a estas últimas de estructura de espacio vectorial en el que se reflejan las propiedades de aquéllos. Este espacio vectorial de las formas lineales se llama «espacio dual» del espacio de vectores del que se partió.

6.9. ESPACIO DUAL

[107]

Sea V un espacio vectorial sobre un cuerpo K. Se llama *forma lineal* en V a toda aplicación lineal del espacio vectorial V en K (espacio vectorial sobre sí mismo). Es decir, la aplicación $\varphi: V \to K$ es una forma lineal si (para cualesquiera $\bar{u}, \bar{v} \in V$ y $\lambda, \mu \in K$) se verifica que:

$$\varphi(\lambda\bar{u} + \mu\bar{v}) = \varphi(\bar{u}) + \mu\varphi(\bar{v})$$

Al espacio vectorial $\mathcal{L}(V, K)$ (véase [096]), de estas formas lineales en V, se le denota por V^* y recibe el nombre de *espacio dual* de V.

OBSERVACIONES

1. De acuerdo con lo dicho en [096], para el espacio dual V^*, es:

$$V^* = \{\varphi: V \to K \,/\, \varphi \text{ lineal}\}$$

$$\left.\begin{matrix}\varphi,\ \psi \in V^* \\ \lambda,\ \mu \in K\end{matrix}\right\} \Rightarrow \quad \lambda\varphi + \mu\psi \in V^*$$

donde

$$\lambda\varphi + \mu\psi : \bar{u} \rightarrowtail (\lambda\varphi + \mu\psi)(\bar{u}) = \lambda\varphi(\bar{u}) + \mu\psi(\bar{u})$$

2. Toda forma lineal $\varphi: V \to K$ es una aplicación sobreyectiva, salvo que sea $\varphi = o$. En efecto: si es $\varphi \neq o$, entonces $\varphi(\vec{u}_0) \neq 0$ para algún $\vec{u}_0 \in V$. Dado entonces cualquier $k \in K$, será

$$\text{para } \vec{u} = \frac{k}{\varphi(\vec{u}_0)}\,\vec{u}_0, \quad \text{es } \varphi(\vec{u}) = \frac{k}{\varphi(\vec{u}_0)}\,\varphi(\vec{u}_0) = k$$

luego las imágenes $\varphi(\vec{u})$ para $\vec{u} \in V$ recorren todo K. En particular, tomando $k = 1$, existe algún $\vec{u}$ tal que $\varphi(\vec{u}) = 1$.

EJEMPLOS

1. Para el espacio vectorial real $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$, de las funciones reales de $\mathbb{R}$ en $\mathbb{R}$, la aplicación

$$\varphi: V \to \mathbb{R}, \quad f \mapsto \varphi(f) = f(1)$$

es una forma lineal.

2. Para el espacio vectorial real $V = \mathbb{R}^3$, una forma lineal lo es la aplicación $\varphi: \mathbb{R}^3 \to \mathbb{R}$ dada por

$$(x, y, z) \mapsto \varphi(x, y. z) = 5x + 8y - 6z$$

El espacio dual V^* estaría constituido por todas las aplicaciones $\varphi: \mathbb{R}^3 \to \mathbb{R}$ del tipo

$$\varphi(x, y, z) = ax + by + cz$$

(para unas ciertas $a, b, c \in \mathbb{R}$). Cada forma lineal queda, pues, caracterizada por la terna (a, b, c), que puede ser considerada como su sistema de coordenadas.

BASE DUAL

En el último ejemplo que acabamos de considerar, se puede apreciar que, operando en la base canónica de $\mathbb{R}^3$, es fácil obtener una base para su dual. Esta base la constituyen las tres formas lineales φ_1, φ_2 y φ_3 siguientes:

$$\varphi_1: (x, y, z) \mapsto \varphi_1(x, y, z) = x$$
$$\varphi_2: (x, y, z) \mapsto \varphi_2(x, y, z) = y$$
$$\varphi_3: (x, y, z) \mapsto \varphi_3(x, y, z) = z$$

Cualquier otra forma lineal se puede poner, en efecto, como combinación de estas tres:

$$\varphi(x, y, z) = ax + by + cz = a\varphi_1(x, y. z) + b\varphi_2(x, y, z) + c\varphi_3(x, y, z)$$

O sea, $\varphi = a\varphi_1 + b\varphi_2 + c\varphi_3$. Además φ_1, φ_2 y φ_3 son independientes y, por ello, constituyen una base, como se había anunciado.

Pues bien, lo aquí dicho se verifica en general, siempre que se considere un espacio V de dimensión finita, como pasamos a comprobar.

[108]

Si V es un espacio vectorial, sobre un cuerpo K, que tiene dimensión n y si $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una base de V, entonces el espacio dual V^* también tiene dimensión n y una de sus bases es $B^* = (\varphi_1, \varphi_2, ..., \varphi_n)$, que se llama *base dual* de B, siendo (para $i = 1, 2, ..., n$)

$$\varphi_i : V \to K \quad , \quad \bar{x} \mapsto \varphi_i(\bar{x}) = x_i$$

(x_i = coordenada i-ésima de $\bar{x}$ en B)

Las coordenadas de una forma lineal $\psi \in V^*$ en la base dual B^* son los escalares $\psi(\bar{e}_1)$, $\psi(\bar{e}_2)$, ..., $\psi(\bar{e}_n)$.

DEMOSTRACION

Obsérvese en primer lugar que las n formas φ_i pertenecen a V^*, es decir, que son lineales. Así ocurre ya que, para cualesquiera $\bar{x}, \bar{y} \in V$ y $\lambda, \mu \in K$ es:

$$\varphi_i(\lambda\bar{x} + \mu\bar{y}) = \text{coordenada } i\text{-ésima de } (\lambda\bar{x} + \mu\bar{y}) =$$

$$= \lambda x_i + \mu y_i = \lambda\varphi_i(\bar{x}) + \mu\varphi_i(\bar{y})$$

Nótese también que $\varphi_i(\bar{e}_j)$ vale 0 si es $i \neq j$ y vale 1 si es $i = j$. Debemos comprobar que $B^* = (\varphi_1, \varphi_2, ..., \varphi_n)$ es un sistema independiente que genera todo V^*. En efecto:

- B^* es linealmente independiente ya que (para $\lambda_1, \lambda_2, ..., \lambda_n \in K$):

$$\sum_{i=1}^{n} \lambda_i\varphi_i = o \;\Rightarrow\; \sum_{i=1}^{n} \lambda_i\varphi_i(\bar{u}) = 0 \quad , \quad \forall\, \bar{u} \in V \;\Rightarrow$$

$$\Rightarrow \sum_{i=1}^{n} \lambda_i\varphi_i(\bar{e}_j) = 0 \;(j = 1, 2, ..., n) \;\Rightarrow\; \lambda_j = 0 \;(j = 1, 2, ..., n)$$

- B^* genera V^*, ya que para cualquiera que sea $\psi \in V^*$ y llamando $x_1, x_2, ..., x_n$ a las coordenadas en B de un $\bar{x} \in V$ cualquiera, se tiene:

$$\psi(\bar{x}) = \psi\left(\sum_{i=1}^{n} x_i\bar{e}_i\right) = \sum_{i=1}^{n} x_i\psi(\bar{e}_i) = \sum_{i=1}^{n} \varphi_i(\bar{x})\psi(\bar{e}_i)$$

y como esto se verifica para todo $x \in V$, se concluye que

$$\psi = \sum_{i=1}^{n} \psi(\bar{e}_i)\varphi_i = \psi(\bar{e}_1)\varphi_1 + \psi(\bar{e}_2)\varphi_2 + \cdots + \psi(\bar{e}_n)\varphi_n$$

luego ψ es combinación lineal de las $\varphi_1, \varphi_2, \ldots, \varphi_n$, como había que comprobar.

Nótese finalmente que la última igualdad informa de que los escalares $\psi(\bar{e}_1), \psi(\bar{e}_2), \ldots, \psi(\bar{e}_n)$ son las coordenadas de ψ en la base B^*.

6.10. APLICACION DUAL O TRASPUESTA

[109]

Sean V y W dos espacios vectoriales, ambos sobre el mismo cuerpo K, y sea V^* y W^* sus respectivos espacios duales. Se llama *aplicación dual o traspuesta* de una aplicación lineal $f: V \to W$ a la siguiente aplicación f^*

$$f^*: W^* \to V^* \quad , \quad \varphi \mapsto f^*(\varphi) = \varphi \circ f$$

Para cualquiera que sea la aplicación lineal f, su aplicación dual f^* también es lineal. Se verifican además las siguientes propiedades:

1. $(\lambda f + \mu g)^* = \lambda f^* + \mu g^* \qquad \forall f, g \in \mathcal{L}(V, W), \quad \forall \lambda, \mu \in K.$
2. $(g \circ f:)^* = f^* \circ g^* \qquad \forall f: \mathcal{L}(V, W), \quad \forall g \in \mathcal{L}(W, U).$

Nótese que f^* transforma cualquier $\varphi: W \to K$ lineal en la aplicación

$$f^*(\varphi): V \xrightarrow{f} W \xrightarrow{\varphi} K \quad , \quad \bar{u} \mapsto f^*(\varphi)(\bar{u}) = (\varphi \circ f)(\bar{u}) = \varphi(f(\bar{u}))$$

que es lineal (ya que resulta de componer aplicaciones lineales) de V en K, esto es, $f^*(\varphi)$ es efectivamente una forma lineal de V^*.

DEMOSTRACION

- La aplicación f^* es lineal ya que, para cualesquiera $\varphi, \psi \in W^*$ y $\lambda, \mu \in K$, se verifica (como ahora veremos) que:

$$f^*(\lambda\varphi + \mu\psi) = \lambda f^*(\varphi) + \mu f^*(\psi)$$

En efecto: para cualquier $\bar{u} \in V$ es:

$$f^*(\lambda\varphi + \mu\psi)(\bar{u}) = ((\lambda\varphi + \mu\psi) \circ f)(\bar{u}) = (\lambda\varphi + \mu\psi)(f(\bar{u})) =$$
$$= \lambda\varphi(f(\bar{u})) + \mu\psi(f(\bar{u})) = \lambda(\varphi \circ f)(\bar{u}) + \mu(\psi \circ f)(\bar{u}) =$$
$$= \lambda f^*(\varphi)(\bar{u}) + \mu f^*(\psi)(\bar{u}) = (\lambda f^*(\varphi) + \mu f^*(\psi))(\bar{u})$$

- Las propiedades 1 y 2 del enunciado son ciertas ya que:

1. Para cualquier $\varphi \in W^*$ y todo $\bar{u} \in V$, es:

$$(\lambda f + \mu g)^*(\varphi)(\bar{u}) = \varphi \circ (\lambda f + \mu g)(\bar{u}) = \varphi((\lambda f + \mu g)(\bar{u})) =$$
$$= \varphi(\lambda f(\bar{u}) + \mu g(\bar{u})) = \lambda\varphi(f(\bar{u})) + \mu\varphi(g(\bar{u})) =$$
$$= \lambda(\varphi \circ f)(\bar{u}) + \mu(\varphi \circ g)(\bar{u}) = \lambda f^*(\varphi)(\bar{u}) + \mu g^*(\varphi)(\bar{u})$$

2. Para cualquiera que sea $\Psi \in U^*$, es

$$\left.\begin{array}{l}(g \circ f)^*(\psi) = \psi \circ (g \circ f) = (\psi \circ g) \circ f \\ (f^* \circ g^*)(\psi) = f^*(g^*(\psi)) = f^*(\psi \circ g) = (\psi \circ g) \circ f\end{array}\right\} \text{(iguales)}$$

APLICACION DUAL Y MATRIZ TRASPUESTA

[110]

Sea V y W dos espacios vectoriales, ambos sobre un mismo cuerpo K, que se suponen de dimensión finita, y sean B y C bases de V y W. Si una aplicación lineal $f: V \to W$ tiene a A como matriz asociada, respecto de las bases B y C, entonces su aplicación dual o traspuesta $f^*: W^* \to V^*$ tiene por matriz asociada, en las bases duales C^* y B^*, a la matriz A', traspuesta de A.

DEMOSTRACION

Sean $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ y $C = (\bar{u}_1, ..., \bar{u}_m)$. Si $A = [a_{ij}]$ es la matriz de f en las bases B y C, entonces

$$f(\bar{e}_j) = \sum_{i=1}^{m} a_{ij}\bar{u}_i \quad (\text{para } j = 1, 2, ..., n)$$

Dada cualquier $\tau \in W^*$, hay que encontrar las coordenadas de $\sigma = f^*(\tau)$ en B^* en función de las coordenadas de τ en C^*. Según lo obtenido en [108], unas y otras coordenadas son $\tau_i = \tau(\bar{u}_i)$ y $\sigma_j = \sigma(\bar{e}_j)$. Por tanto, se tiene:

$$[f^*(\tau)]_j = \sigma_j = \sigma(\bar{e}_j) = f^*(\tau)(\bar{e}_j) = (\tau \circ f)(\bar{e}_j) = \tau(f(\bar{e}_j)) = \tau\left(\sum_{i=1}^{m} a_{ij}\bar{u}_i\right) =$$

$$= \sum_{i=1}^{m} a_{ij}\tau(\bar{u}_i) = \sum_{i=1}^{m} a_{ij}\tau_i \quad (\text{para } j = 1, 2, ..., n)$$

Así, pues, el elemento de lugar ji de la matriz asociada a f^* es igual al elemento de lugar ij de la matriz de f; esto es, la matriz de f^* es la traspuesta de la matriz de f.

ESPACIO BIDUAL

[111]

Dado un espacio vectorial V, sobre un cuerpo K, se llama *espacio bidual* de V al espacio dual de V^*, al que se representa por V^{**}.
Para cada $\bar{u} \in V$, la aplicación $F_{\bar{u}}: V^* \to K$, $F_{\bar{u}}(\varphi) = \varphi(\bar{u})$, es un elemento de V^{**}. La aplicación $F: V \to V^{**}$, $F(\bar{u}) = F_{\bar{u}}$ es lineal inyectiva. Si V es de dimensión finita, F es un isomorfismo, que permite identificar V con su bidual V^{**}, que recibe el nombre de isomorfismo canónico.

DEMOSTRACION

- Para cualquier $\bar{u} \in V$, la aplicación $F_{\bar{u}}$ es lineal (luego pertenece a V^{**}) ya que, para $\varphi, \Psi \in V^*$ y $\lambda, \mu \in K$, es:

$$F_{\bar{u}}(\lambda\varphi + \mu\psi) = (\lambda\varphi + \mu\psi)(\bar{u}) = \lambda\varphi(\bar{u}) + \mu\psi(\bar{u}) = \lambda F_{\bar{u}}(\varphi) + \mu F_{\bar{u}}(\psi)$$

- La aplicación F es lineal ya que, para $\bar{u}, \bar{v} \in V$ y $\lambda, \mu \in K$, es

$$F(\lambda\bar{u} + \mu\bar{v}) = \lambda F(\bar{u}) + \mu F(\bar{v})$$

puesto que para cualquiera que sea $\varphi \in V^*$, se verifica que:

$$F(\lambda\bar{u} + \mu\bar{v})(\varphi) = F_{\lambda\bar{u}+\mu\bar{v}}(\varphi) = \varphi(\lambda\bar{u} + \mu\bar{v}) = \lambda\varphi(\bar{u}) + \mu\varphi(\bar{v}) =$$
$$= \lambda F_{\bar{u}}(\varphi) + \mu F_{\bar{v}}(\varphi) = (\lambda F_{\bar{u}} + \mu F_{\bar{v}})(\varphi) = (\lambda F(\bar{u}) + \mu F(\bar{v}))(\varphi)$$

- La aplicación lineal F es inyectiva ya que su núcleo es nulo, pues está formado por las $\bar{u} \in V$ tales que $F_{\bar{u}} = o$, esto es, tales que $F_{\bar{u}}(\varphi) = 0$ para todo $\varphi \in V^*$, o sea, tales que $\varphi(\bar{u}) = 0$ para toda forma lineal $\varphi \in V^*$, lo cual sólo ocurre para $\bar{u} = \bar{o}$.
- Si V tiene dimensión finita, $\dim V = n$, entonces V^* también tiene dimensión n (véase [108]). Como $V^{**} = (V^*)^*$, resulta que V^{**} tiene la misma dimensión que V^*, es decir V^{**} tiene dimensión n. Este resultado conduce a que la aplicación lineal $F: V \to V^{**}$ que es inyectiva sea, ahora, también sobreyectiva, es decir, a que F sea un isomorfismo.

Ejercicios y problemas a la parte II

ENUNCIADOS

II.1. Sean V_1 y V_2 dos espacios vectoriales (con el mismo cuerpo de escalares) y considérese el producto cartesiano $V_1 \times V_2$, en el que se definen la suma y el producto por un escalar mediante:

$$(\vec{v}_1, \vec{v}_2) + (\vec{v}_1', \vec{v}_2') = (\vec{v}_1 + \vec{v}_1', \vec{v}_2 + \vec{v}_2')$$
$$\lambda(\vec{v}_1, \vec{v}_2) = (\lambda\vec{v}_1, \lambda\vec{v}_2)$$

Pruébese que con estas operaciones el conjunto $V_1 \times V_2$ es un espacio vectorial. Si V_1 y V_2 tienen dimensión finita, compruébese que:

$$\dim (V_1 \times V_2) = \dim V_1 + \dim V_2$$

II.2. Compruébese que el conjunto $\mathscr{C}(\mathbb{R}, \mathbb{R})$, de las funciones continuas de $\mathbb{R}$ en $\mathbb{R}$, es un espacio vectorial real respecto de las operaciones usuales. Se consideran las siguientes funciones de $\mathscr{C}(\mathbb{R}, \mathbb{R})$:

$$\text{sen } x,\ 1,\ x,\ x^2,\ \ldots,\ x^n \quad (n \in \mathbb{N} \text{ dado})$$

Analícese si la primera es una combinación lineal de las demás.

II.3. Considérese el espacio vectorial $V = \mathcal{M}_{2\times2}$, de las matrices cuadradas de tamaño 2×2, y sea $S = (M_1, M_2, M_3, M_4)$ el sistema formado por las matrices:

$$M_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$M_3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

a) Comprobar que S es una base de V.
b) Hallar las coordenadas x_1, x_2, x_3, x_4 en la base S de una matriz genérica M de V:

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

II.4. En el espacio vectorial V de los polinomios de grado menor o igual que 4, se consideran los polinomios:

$$p_1(x) = 3 - 2x + x^2 + 4x^3 + x^4$$
$$p_2(x) = 4 - x + x^2 + 6x^3 - 2x^4$$
$$p_3(x) = 7 - 8x + 3x^2 + ax^3 + bx^4$$

($a, b \in \mathbb{R}$ fijos). Hallar a y b para que el subespacio que engendran $p_1(x)$, $p_2(x)$ y $p_3(x)$ tenga dimensión 2. Hallar una base cualquiera de este subespacio y determinar las coordenadas en ellas de los tres polinomios dados.

II.5. Hallar el valor que hay que asignar al parámetro α para que las siguientes matrices no formen base del espacio vectorial $\mathcal{M}_{2\times2}$, de las matrices cuadradas de tamaño 2×2:

$$\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 6 & 5 \\ 4 & 2 \end{bmatrix}, \begin{bmatrix} 5 & 4 \\ 4 & \alpha \end{bmatrix}$$

II.6. En el espacio vectorial de las matrices reales y simétricas de tamaño 3×3, se consideran las matrices:

$$M_1 = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 4 \\ -1 & 4 & -2 \end{bmatrix} \quad M_2 = \begin{bmatrix} 2 & 4 & -2 \\ 4 & 6 & 8 \\ -2 & 8 & -4 \end{bmatrix}$$

$$M_3 = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 6 \\ 2 & 6 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 1 & 4 & 5 \\ 4 & 1 & 8 \\ 5 & 8 & 2 \end{bmatrix}$$

$$M_5 = \begin{bmatrix} 2 & 7 & 3 \\ 7 & 3 & 9 \\ 3 & 9 & 1 \end{bmatrix}$$

Hallar la dimensión y una base del subespacio que engendran estas 5 matrices. Obtener las coordenadas de todas ellas en la base elegida.

II.7. En el espacio vectorial $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$, de las funciones de $\mathbb{R}$ en $\mathbb{R}$, se consideran el sistema de funciones:

$$S = (1, \text{sen } x, \cos x, \text{sen } 2x, \cos 2x)$$

a) Comprobar que S es linealmente independiente.

b) Hallar una base B del subespacio que engendran las funciones:

$$f_1(x) = 1 - 2 \operatorname{sen} x + 3 \cos x - \operatorname{sen} 2x$$
$$f_2(x) = \operatorname{sen} x + \cos x - 2 \operatorname{sen} 2x - \cos 2x$$
$$f_3(x) = 2 - \cos x + \operatorname{sen} 2x + 3 \cos 2x$$
$$f_4(x) = 1 + 4 \operatorname{sen} x - 2 \cos x - 2 \operatorname{sen} 2x + \cos 2x$$
$$f_5(x) = 4 + \operatorname{sen} x - \cos x + 5 \cos 2x$$

c) Completar la base B hasta obtener una base del subespacio engendrado por S.

II.8. En el espacio vectorial V de los polinomios reales de grado menor o igual que n (con una sola indeterminada x), se considera la base usual $B = (1, x, x^2, \ldots, x^n)$. Compruébese que $B' = (1, x - a, (x-a)^2, \ldots, (x-a)^n)$, donde $a \in \mathbb{R}$ es dado, es una base de V y hállese la matriz del cambio de coordenadas que resulta de cambiar la base B por la B'.

II.9. En el espacio vectorial $V = \mathcal{M}_{2\times 2}$, de las matrices reales de tamaño 2×2, se consideran los sistemas $S = (M_1, M_2, M_3, M_4)$ y $T = (N_1, N_2, N_3, N_4, N_5)$, donde:

$$M_1 = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \qquad M_2 = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}$$
$$M_3 = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \qquad M_4 = \begin{bmatrix} 0 & -1 \\ 3 & -2 \end{bmatrix}$$
$$N_1 = \begin{bmatrix} -2 & 1 \\ 3 & -2 \end{bmatrix} \qquad N_2 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$
$$N_3 = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \qquad N_4 = \begin{bmatrix} 0 & 4 \\ 2 & 1 \end{bmatrix}$$
$$N_5 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

Comprobar si S y T son sistemas equivalentes. Hallar bases de los subespacios que engendran S y T.

II.10. Sea V el espacio vectorial de los polinomios de grado menor o igual que n (dado). Sea $p(x)$ un polinomio cualquiera de V y considérese el sistema

$$B = (p(x), p'(x), p''(x), \ldots, p^{(n)}(x))$$

Analizar si B es o no una base de V, en función de quién sea $p(x)$.

II.11. Sea A una matriz cuadrada de tamaño $n \times n$; sean $C_1, C_2, \ldots, C_p$ matrices columna de tamaño $n \times 1$. Pruébese que:

a) Si los vectores columna $AC_1, AC_2, \ldots, AC_p$ son linealmente independientes, entonces también lo son los $C_1, C_2, \ldots, C_p$.

b) Si A es regular y $C_1, C_2, \ldots, C_p$ son linealmente independientes, entonces también lo son $AC_1, AC_2, \ldots, AC_p$.

II.12. En el espacio vectorial $\mathcal{F}(\mathbb{R}, \mathbb{R})$, de las funciones de $\mathbb{R}$ en $\mathbb{R}$. se consideran las funciones:

$$f_1(x) = e^{a_1 x}, f_2(x) = e^{a_2 x}, \ldots, f_n(x) = e^{a_n x}$$

Compruébese que estas n funciones son linealmente independientes si y sólo si los números reales $a_1, a_2, \ldots, a_n$ son todos distintos.

II.13. Sean $S = (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p)$ y $S' = (\vec{u}'_1, \vec{u}'_2, \ldots, \vec{u}'_p)$ dos sistemas de p vectores, de un cierto espacio vectorial V. Demuéstrese que:

$$\operatorname{rang}(\vec{u}_1 + \vec{u}'_1, \vec{u}_2 + \vec{u}'_2, \ldots, \vec{u}_p + \vec{u}'_p) \leqslant \leqslant \operatorname{rang} S + \operatorname{rang} S'$$

Si A y A' son dos matrices de igual tamaño, compruébese que:

$$\operatorname{rang}(A + A') \leqslant \operatorname{rang} A + \operatorname{rang} A'$$

II.14. Demostrar que, si U, V y W son subespacios vectoriales de un cierto espacio, entonces:

a) $(U \cap V) + (U \cap W) \subset U \cap (V + W)$

b) $(U + V) \cap (U + W) \supset U + (V \cap W)$

Determinar algún ejemplo en el que los signos de inclusión de *a*) y *b*) no lo sean de igualdad.

II.15. Considérense los sistemas de polinomios $S = (p_1(x), p_2(x), p_3(x))$ y $T = (q_1(x), q_2(x), q_3(x))$, donde:

$$p_1(x) = 1 + 2x + 5x^2 + 3x^3 + 2x^4$$
$$p_2(x) = 3 + x + 5x^2 - 6x^3 + 6x^4$$
$$p_3(x) = 1 + x + 3x^2 + 2x^4$$

$$q_1(x) = 2 + x + 4x^2 - 3x^3 + 4x^4$$
$$q_2(x) = 3 + x + 3x^2 - 2x^3 + 2x^4$$
$$q_3(x) = 9 + 2x + 3x^2 - x^3 - 2x^4$$

Sean U y V los subespacios, del espacio vectorial de los polinomios reales, que engendran S y T. Hallar la dimensión y una base de cada uno de los subespacios $U + V$ y $U \cap V$.

II.16. Sean $U_1, U_2, ..., U_p$ subespacios de un cierto espacio vectorial V. Compruébese que la suma $U_1 + U_2 + \cdots + U_p$ es directa si y sólo si para cualquier $i = 1, 2, ..., p$ se verifica que:

$$U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_p) = O$$

II.17. Sean U_1 y U_2 los siguientes subespacios del espacio vectorial real $\mathbb{R}^n$:

$$U_1 = \{(x_1, x_2, ..., x_n) \in \mathbb{R}^n / x_1 + x_2 + \cdots + x_n = 0\}$$
$$U_2 = \{(x_1, x_2, ..., x_n) \in \mathbb{R}^n / x_1 = x_2 = \cdots = x_n\}$$

Analizar si U_1 y U_2 son subespacios suplementarios de $\mathbb{R}^n$ obteniendo la descomposición de cualquier $\bar{u} \in \mathbb{R}$ en suma $\bar{u} = \bar{u}_1 + \bar{u}_2$, con $\bar{u}_1 \in U_1$ y $\bar{u}_2 \in U_2$.

II.18. Sea $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ el espacio vectorial de las funciones de $\mathbb{R}$ en $\mathbb{R}$ y considérense los conjuntos:

$$U_1 = \{f \in V / f(1) = f(-1) = 0\}$$
$$U_2 = \{f \in V / f \text{ es afín}\}$$

(la función f se dice afín si es $f(x) = a + bx$ para ciertos a y b fijos). Pruébese que U_1 y U_2 son subespacios suplementarios de V.

II.19. Sea $V = \mathcal{C}([0,1]), \mathbb{R})$ el espacio vectorial de las funciones continuas de $[0, 1]$ en $\mathbb{R}$. Se pide:

a) Dado un número natural n y si $f \in V$ es una función que toma, al menos, n valores distintos, analizar si las n funciones $f(x)$, $f(x)^2$, $f(x)^3$, ..., $f(x)^n$ forman un sistema linealmente independiente de V.

b) Analizar si U_1 y U_2 son subespacios suplementarios de V, siendo:

$$U_1 = \left\{ f \in V \Big/ \int_0^1 f(x)\,dx = 0 \right\}$$

y

$$U_2 = \{f \in V / f \text{ es constante}\}$$

II.20. Sean U, V, U_1, U_2, V_1 y V_2 subespacios de un cierto espacio vectorial. Si la suma $U + V$ es directa y si $U = U_1 \oplus U_2$ y $V = V_1 \oplus V_2$, compruébese que la suma $U_1 + U_2 + V_1 + V_2$ es directa.

II.21. Si A y B son dos matrices cuadradas del mismo tamaño simétricas, compruébese que AB es simétrica si y sólo si $AB = BA$.

II.22. Sea A una matriz $m \times p$ y sea B una matriz $p \times n$. Pruébese que:

1. rang $(AB) \leq$ rang A
2. rang $(AB) \leq$ rang B

II.23. Sean A y N dos matrices cuadradas de igual tamaño tales que A es invertible, N es mihilpotente (o sea, $N^2 = O$) y N conmuta con A^{-1}. Pruébese que $A + N$ es inversible hallando su matriz inversa. (Indicación: recúrrase a la igualdad $(I + M)^{-1} = I - M + M^2 - M^3 ...$)

II.24. Analícese si el conjunto $\mathcal{M}$ formado por todas las matrices de la forma:

$$A = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \quad (a, b, c \in \mathbb{R})$$

es o no un grupo multiplicativo.

II.25. Sean V y W dos espacios vectoriales reales; sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3, \bar{e}_4)$ una base de V y sea $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ una base de W. Si $f: V \to W$ es la aplicación lineal definida por:

$$f(\bar{e}_1) = \bar{u}_1 + \bar{u}_2 - 4\bar{u}_3 \qquad f(\bar{e}_3) = 3\bar{u}_1 + \bar{u}_2$$
$$f(\bar{e}_2) = 2\bar{u}_1 + \bar{u}_2 - 2\bar{u}_3 \qquad f(\bar{e}_4) = \bar{u}_1 + 2\bar{u}_3$$

se pide hallar:

a) Ecuaciones de f en las bases dadas.

b) Una base de Nuc (f) y otra de Im (f).

II.26. Sea dado un espacio vectorial V de dimensión finita y considérese un endomorfismo $f: V \to V$. Pruébese que, si f y $f \circ f$ tienen igual rango, entonces Nuc $(f) \cap$ Im $(f) = O$.

II.27. Sea $f: \mathbb{R}^3 \to \mathbb{R}^2$ la aplicación lineal que respecto de las bases canónicas tiene asociada a la siguiente matriz A:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 1 \end{bmatrix}$$

a) Hallar la matriz D canónica de equivalencia de A.
b) Hallar unas bases $B = (\bar{e}_1, \bar{e}_2, \bar{e}_3)$ de $\mathbb{R}^3$ y $C = (\bar{u}_1, \bar{u}_2)$ de $\mathbb{R}^2$ respecto de las cuales la matriz de f sea la canónica.
c) Analizar si existe alguna aplicación lineal g tal que alguna de las $f \circ g$ o $g \circ f$ sea la identidad.

II.28. Sea $i: \mathbb{R}^3 \to \mathbb{R}^3$ la aplicación identidad y llamemos $B = (\bar{e}_1, \bar{e}_2, \bar{e}_3)$ a la base canónica de $\mathbb{R}^3$. Se pide:

a) Una base B' de $\mathbb{R}^3$ tal que al referir la aplicación i a B' (base del espacio $\mathbb{R}^3$ de partida) y a B (base del espacio $\mathbb{R}^3$ de llegada), la matriz asociada sea la A que figura al final del enunciado.
b) Una base B'' de $\mathbb{R}^3$ tal que al referir la aplicación i a B (base del espacio $\mathbb{R}^3$ de partida) y a B'' (base del espacio $\mathbb{R}^3$ de llegada), la matriz asociada sea A:

$$A = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

II.29. Sea V el conjunto formado por todas aquellas matrices $A = [a_{ij}]$ cuadradas de tamaño $n \times n$ tales que las sumas de los elementos de cada una de sus filas y de cada una de sus columnas son, todas, iguales; llamaremos a este número $S(A)$:

$$S(A) = a_{i1} + a_{i2} + \cdots + a_{in} = a_{1j} + a_{2j} + \cdots + a_{nj} \quad (\forall\, i, j = 1, 2, \ldots, n)$$

a) Analizar si V es un espacio vectorial (con las operaciones usuales) y si $S: V \to \mathbb{R}$, $A \mapsto S(A)$, es una forma lineal.
b) Llamando J a la matriz $n \times n$ que tiene todos sus elementos iguales a la unidad, pruébese que:

$$[A \in V] \Leftrightarrow [\exists\, \alpha \in \mathbb{R},\ AJ = JA = \alpha J]$$

Hállese α en función de A.
c) Si $A \in V$ es regular, compruébese que $S(A) \neq 0$ y que $A^{-1} \in V$. Hallar $S(A^{-1})$ en función de $S(A)$.

II.30. Sea $\mathcal{M}$ el conjunto de las matrices reales cuadradas y que conmutan con una cierta matriz dada. Compruébese que $\mathcal{M}$ es un anillo respecto de las operaciones usuales. Suponiendo a partir de ahora que las matrices son de tamaño 2×2 y si la matriz dada es:

$$M = \begin{bmatrix} 1 & h \\ -1 & 1 \end{bmatrix} \quad (h \text{ parámetro dado})$$

se pide:

1. Expresión general de las matrices de $\mathcal{M}$.
2. Estudiar si en $\mathcal{M}$ hay divisores de cero.
3. Estudiar si $\mathcal{M}$ es un cuerpo.

II.31. Sea $\mathcal{M}$ el conjunto formado por las matrices reales de tamaño $n \times n$ que son de la forma:

$$A(x_1, x_2, \ldots, x_n) = \begin{bmatrix} x_1 & 0 & 0 & \vdots & \vdots & 0 \\ x_2 & x_1 & 0 & \vdots & \vdots & 0 \\ x_3 & x_2 & x_1 & \vdots & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & x_{n-1} & x_{n-2} & \vdots & \vdots & x_1 \end{bmatrix}$$

donde $x_1, x_2, \ldots, x_n$ recorren $\mathbb{R}$. Se pide:

1. Expresar $A(x_1, x_2, \ldots, x_n)$ en función de $I = A(1, 0, \ldots 0)$, de $E = A(0, 1, \ldots, 0)$ y de las potencias de E.
2. Comprobar que $\mathcal{M}$ es un espacio vectorial y hallar una de sus bases.
3. Hallar la matriz inversa de $A(1, 2, 3, \ldots, n)$.

II.32. Sea $f: \mathbb{R}^4 \to \mathbb{R}^4$ un endomorfismo del que se sabe que:

$$f(1, 1, 0, 0) = (0, 1, 0, -1)$$

$$f(1, 0, 1, 0) = (1, 1, 1, 0)$$

Hallar la matriz asociada a f, respecto de las bases canónicas, en cada uno de los siguientes supuestos:

1. Nuc (f) = Im (f).
2. $f \circ f = i$ (identidad).
3. $f \circ f = f$.

II.33. Sean $(\bar{e}_1, \bar{e}_2, \bar{e}_3, \bar{e}_4)$ y $(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4, \bar{u}_5)$ las bases canónicas de $\mathbb{R}^4$ y $\mathbb{R}^5$. Considérese la aplicación lineal $f: \mathbb{R}^4 \to \mathbb{R}^5$ definida por:

$$f(\bar{e}_1) = \bar{u}_1 + \bar{u}_3 - \bar{u}_5$$
$$f(\bar{e}_2) = 2\bar{u}_2 - \bar{u}_3 + \bar{u}_4 + 2\bar{u}_5$$
$$f(\bar{e}_3) = \bar{u}_1 + 2\bar{u}_2 + \bar{u}_4 + \bar{u}_5$$
$$f(\bar{e}_4) = \bar{u}_1 + 4\bar{u}_2 - \bar{u}_3 + 2\bar{u}_4 + 3\bar{u}_5$$

Se pide:

1. Ecuaciones de f en las bases canónicas de $\mathbb{R}^4$ y $\mathbb{R}^5$.
2. Ecuaciones de f en la base $(\bar{e}_1 + \bar{e}_2,\ \bar{e}_2 - \bar{e}_3,\ \bar{e}_1 + \bar{e}_3 + \bar{e}_4,\ \bar{e}_3 - \bar{e}_4)$ de $\mathbb{R}^4$ y la canónica de $\mathbb{R}^5$.
3. Hallar bases de Nuc (f) y de Im (f).
4. Hallar un subespacio V de $\mathbb{R}^4$ tal que la restricción de f a V sea inyectiva y tenga la misma imagen que f.

II.34. Sea V el espacio vectorial de los polinomios reales de grados menor o igual que 2; sean:

$$p(x) = 1 + x + x^2 \quad q(x) = 1 + 2x^2 \quad r(x) = x + x^2$$
$$\bar{a} = (2, 0, 1) \quad \bar{b} = (3, 1, 0) \quad \bar{c} = (1, -2, 3)$$

Considérese la aplicación lineal $f: V \to \mathbb{R}^3$ definida por:

$$f(p(x)) = \bar{a} \quad f(q(x)) = \bar{b} \quad f(r(x)) = \bar{c}$$

1. Hallar la matriz de f respecto de las bases $(1, x, x^2)$ de V y la canónica de $\mathbb{R}^3$.
2. Hallar la matriz de f respecto de las bases $(p(x), q(x), r(x))$ de V y $(\bar{a}, \bar{b}, \bar{c})$ de $\mathbb{R}^3$.
3. Hallar una base B de V, tal que respecto de ella y de la base canónica de $\mathbb{R}^3$ la matriz de f sea la unidad I (compruébese previamente que ello es posible).

II.35. Dadas las matrices de igual tamaño:

$$A = \begin{bmatrix} 1 & 5 & -1 & 2 \\ 2 & 1 & 4 & 1 \\ 1 & 2 & 1 & a \end{bmatrix} \quad B = \begin{bmatrix} 2 & 0 & 1 & 2 \\ 1 & -1 & 2 & 3 \\ 0 & 2 & -3 & -4 \end{bmatrix}$$

donde a es el número real para el que A y B son equivalentes:

1. Hallar a.
2. Hallar dos matrices regulares P y Q tales que $B = Q^{-1}AP$.
3. Si es $f: \mathbb{R}^4 \to \mathbb{R}^3$ la aplicación lineal asociada a la matriz A respecto a las bases canónicas, hallar una base B de $\mathbb{R}^4$ y otra base C de $\mathbb{R}^3$ respecto de las que la matriz asociada a f sea B.

II.36. Sea V el espacio vectorial de los polinomios reales de grado menor o igual que 2 y considérese la aplicación lineal $f: \mathbb{R}^3 \to V$ en la que (para ciertos $\alpha, \beta \in \mathbb{R}$):

$$f(1, 1, 1) = 2\beta + \alpha x \quad , \quad f(0, -1, 1) = \alpha x + \beta x^2$$
$$f(0, 0, 1) = \beta + (\alpha - 1)x$$

1. Hallar α y β para que f no sea inyectiva.
2. Hallar bases de Nuc (f) y de Im (f), en función de α y β.
3. Hallar el subespacio imagen por la aplicación f, del $U = \{(a, b, c) \in \mathbb{R}^3 / a + c = b + c = 0\}$, según los valores de α y β.

II.37. Sean $f: \mathbb{R}^3 \to \mathbb{R}^4$ y $g: \mathbb{R}^4 \to \mathbb{R}^3$ dos aplicaciones lineales tales que:

a) $h = g \circ f$ es la proyección sobre el subespacio $\{(y_1, y_2, y_3) \in \mathbb{R}^3 / y_2 = y_3\}$ paralelamente al vector $(1, -1, 0)$.

b) $k = f \circ g$ es tal que su núcleo es Nuc $(k) = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 / x_3 = 0,\ x_1 + x_2 = x_4\}$ y además $k(1, 0, 0, 0) = (1, 1, 2, 1)$ y $k(0, 0, 1, 0) = (1, 0, 1, 1)$.

En las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^4$, se pide:

1. La matriz de h y sendas bases de Im (h) y de Nuc (h).
2. Matriz de k y sendas bases de Im (k) y Nuc (k).
3. Sabiendo que g es una aplicación sobreyectiva, hallar el núcleo y la imagen de f.

II.38. Sea $f: V \to W$ una aplicación lineal de la que se conoce Im (f) y un subespacio $U \subset V$ de Nuc (f). Si es $\varphi: U \to W$ la restricción de f a U, hallar el Nuc (φ) y la Im (φ).

II.39. Sea V el espacio vectorial real de las funciones de $\mathbb{R}$ en $\mathbb{R}$, con las operaciones usuales. Considérese las aplicaciones $F: \mathbb{R}^3 \to V$ y $G: V \to \mathbb{R}^2$ definidas por:

$$F(a, b, c) = a \operatorname{sen}^2 x + b \cos^2 x + c$$
$$G(f(x)) = (f(0), f(-\pi/2))$$

1. Comprobar que F y G son lineales y hallar sus núcleos y sus imágenes.
2. Si $U \subset V$ es el subespacio de las funciones constantes, hallar $F^{-1}(U)$.
3. Hallar bases de $\mathbb{R}^3$ y de $F(\mathbb{R}^3)$ respecto de las que la matriz de F sea la canónica (de equivalencia).
4. Hallar la ecuación y el núcleo de la aplicación $G \circ F$.

II.40. Sea $\varphi: V \to V$ un endomorfismo del espacio vectorial V de dimensión n. Compruébese que los endomorfismos $f: V \to V$ tales que $\varphi \circ f = o$ forman en conjunto $\mathbb{V}$ que es un espacio vectorial; hallar la dimensión de éste.

II.41. Sea $f: V \to W$ una aplicación lineal cualquiera entre espacios vectoriales de igual dimensión (finita). Compruébese que f se puede expresar como suma de dos isomorfismos.

II.42. Sea $f: V \to W$ una aplicación lineal entre espacios vectoriales de dimensión finita; sea dim $V = n$ y dim $W = m$. Pruébese que f es inyectiva si y sólo si existe una aplicación lineal $g: W \to V$ tal que $g \circ f = i$ (identidad en V).

II.43. Sea $f: V \to W$ una aplicación lineal entre espacios vectoriales de dimensión finita; sea dim $V = n$ y dim $W = m$. Pruébese que f es sobreyectiva si y sólo si existe una aplicación lineal $g: W \to V$ tal que $f \circ g = i$ (identidad en W).

II.44. Sea $f: V \to V$ un endomorfismo tal que $i + af + bf \circ f = o$, donde i es la identidad en V y a y b son dos escalares dados. Comprobar que f es un automorfismo hallando su inverso.

II.45. Sea $f: V \to V$ un endomorfismo en V con dim $V = n$. Compruébese que Nuc (f) = Im (f) si y sólo si n es par, rang $f = n/2$ y $f \circ f = o$.

II.46. Sea V un espacio vectorial real de dimensión 3 y sea $B = (\bar{e}_1, \bar{e}_2, \bar{e}_3)$ una base de V. Se considera la forma lineal $\varphi: V \to \mathbb{R}$ que satisface a:

$$\varphi(\bar{e}_1 - \bar{e}_2) = 5 \ ; \ \varphi(\bar{e}_2 - \bar{e}_3) = 3 \ ; \ \varphi(\bar{e}_1 + \bar{e}_3) = 4$$

Hallar:

a) La ecuación de φ en la base B.

b) Coordenadas de φ en la base B^*, dual de B.

II.47. Sea $B = (\bar{e}_1, \bar{e}_2, \bar{e}_3)$ la base canónica de $\mathbb{R}^3$ y considérense tres formas lineales φ, ψ y η, en $\mathbb{R}^3$, de las que se sabe:

- $\varphi(\bar{e}_1) = 1$, $\varphi(\bar{e}_2) = a$, $\varphi(\bar{e}_3) = 0$ ($a \in \mathbb{R}$ dado).
- $\psi(\bar{e}_1 + \bar{e}_2) = 3$, $\psi(\bar{e}_1 + \bar{e}_2 - 2\bar{e}_3) = 1$, $\psi(2\bar{e}_1 + \bar{e}_2 - \bar{e}_3) = 2$.
- Nuc $(\eta) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 / x_1 + x_2 + bx_3 = 0\}$, $\eta(\bar{e}_1) = 2$ ($b \in \mathbb{R}$ dado).

1. Hallar las ecuaciones de φ, ψ y η en la base B.
2. Hallar las coordenadas de φ, ψ y η en la base B^*, dual de B.
3. Hallar la relación que deben guardar a y b para que (φ, ψ, η) sea un sistema linealmente dependiente de V^*.

II.48. En el espacio vectorial $\mathbb{R}^3$ se considera la siguiente base $B = (\bar{u}_1, \bar{u}_2, \bar{u}_3)$:

$$\bar{u}_1 = (1, -1, 3) \ , \ \bar{u}_2 = (0, 1, -1) \ , \ \bar{u}_3 = (0, 3, -2)$$

Hallar la base dual $B^* = (\varphi_1, \varphi_2, \varphi_3)$ de la base B.

II.49. Sea V el espacio vectorial de los polinomios reales de grado menor o igual que 1. Sean $\varphi_1: V \to \mathbb{R}$ y $\varphi_2: V \to \mathbb{R}$ las formas lineales:

$$\varphi_1(p(x)) = \int_0^1 p(x)\,dx \quad ; \quad \varphi_2(p(x)) = \int_0^2 p(x)\,dx$$

Comprobar que (φ_1, φ_2) es una base del espacio V^*, dual de V, y hallar la base $(p_1(x), p_2(x))$ de V que tiene a (φ_1, φ_2) por dual.

II.50. Sea V un espacio vectorial sobre un cuerpo K y sea $f: V \to K$ una forma lineal. Compruébese que el subespacio U = Nuc (f) es máximal de V, o sea, es suplementario de un subespacio unidimensional de V.

II.51. Sea V un espacio vectorial, sobre un cuerpo K, y sea U un subespacio máximal de V, esto es, U es suplementario de un subespacio unidimensional de V. Pruébese que existe una forma lineal $f: V \to K$ cuyo núcleo es U.

SOLUCIONES

I.1. Indicación: si $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ es una base de V_1 y $(\bar{v}_1, \bar{v}_2, ..., \bar{v}_m)$ es una base de V_2, compruébese que los siguientes vectores forman una base de $V_1 \times V_2$

$$(\bar{u}_1, \bar{o}), (\bar{u}_2, \bar{o}), ..., (\bar{u}_n, \bar{o}), (\bar{o}, \bar{v}_1), (\bar{o}, \bar{v}_2), ..., (\bar{o}, \bar{v}_m)$$

II.2. $\mathscr{C}(\mathbb{R}, \mathbb{R})$ es subespacio de $\mathscr{F}(\mathbb{R}, \mathbb{R})$. La función $\operatorname{sen} x$ no es combinación lineal de $1, x, x^2, ..., x^n$ pues aquélla se anula para infinitos valores de x.

II.3. $x_1 = a - b + c + d$, $x_2 = -a + b + c$, $x_3 = a - c$, $x_4 = c$.

II.4. $a = 8$, $b = 9$; $(p_1(x), p_2(x))$ es base, $p_3(x) = 5p_1(x) - 2p_2(x)$.

II.5. $\alpha = 3$.

II.6. La dimensión es 3; una base es la $B = (M_1, M_3, M_5)$; $M_2 = 2M_1$, $M_4 = -M_1 + 2M_3$.

II.7. La relación $a + b \operatorname{sen} x + c \cos x + d \operatorname{sen} 2x + e \cos 2x = 0$ para todo $x \in \mathbb{R}$ sólo tiene la solución $a = b = c = d = e = 0$ (tómense para x los valores 0, $\pi/2$, π, $3\pi/2$, $\pi/4$, por ejemplo). $B = (f_1(x), f_2(x), f_3(x))$. Al añadir a B las funciones $\operatorname{sen} 2x$ y $\cos 2x$ se obtiene una base de $\mathscr{V}(S)$.

II.8.
$$\left[\begin{array}{cccc:c} 1 & -a & a^2 & -a^3 & (-1)^n a^n \\ 0 & 1 & 2a & 3a^2 & -(-1)^n \binom{n}{1} a^{n-1} \\ 0 & 0 & 1 & -3a & (-1)^n \binom{n}{2} a^{n-2} \\ 0 & 0 & 0 & 1 & -(-1)^n \binom{n}{3} a^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Nótese que esta matriz es regular para todo $a \in \mathbb{R}$.

II.9. Recurrir a la base usual. Los dos sistemas son equivalentes y engendran el mismo espacio; una base de éste es la $B = (E_1, E_2, E_3)$, donde:

$$E_1 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \quad E_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad E_3 = \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}$$

II.10. El sistema B es base si y sólo si $p(x)$ es de grado n.

II.11. Indicación $\Sigma \lambda_i(AC_i) = A(\Sigma \lambda C_i)$. Recurrir al contrarrecíproco. Si A es regular, de ser $\Sigma \lambda_i(AC_i) = O$, sería $\Sigma \lambda C_i = A^{-1} \cdot O = O$.

II.12. Recurrir al método de inducción. La relación $\Sigma \lambda_i f_i(x) \equiv 0$ $(i = 1, 2, ..., n)$, si se la divide por $f_n(t)$ y, luego, se deriva conduce a la correspondiente relación para $n - 1$ funciones.

II.13. Sean $U = \mathscr{V}(S)$ y $U' = \mathscr{V}(S')$. Los vectores $\bar{u}_i + \bar{u}_i'$ pertenecen a $U + U'$, luego:

$$\operatorname{rang}(\bar{u}_i + \bar{u}_i') \leqslant \dim(U + U') \leqslant$$
$$\leqslant \dim U + \dim U' = \operatorname{rang} S + \operatorname{rang} S'$$

Para el caso de las matrices, recurrir a sus vectores columna (o fila) y aplicar el resultado anterior.

II.14. *a*) Téngase en cuenta que $V \subset V + W$ y que $W \subset V + W$; *b*) téngase en cuenta que $V \supset V \cap W$ y que $W \supset V \cap W$. Contraejemplo: en $\mathbb{R}^2$ tomar:

$$U = \{(t, t) / t \in \mathbb{R}\} \quad V = \{(x, 0) / x \in \mathbb{R}\}$$

y

$$W = \{(0, y) / y \in \mathbb{R}\}$$

II.15. La dimensión de $U + V$ es 3 y una base suya es la $(p_1(x), p_2(x), q_1(x))$. La dimensión de $U \cap V$ es 1 y una base suya es la $(2 + x + 4x^2 - 3x^3 + 4x^4)$.

II.16. Indicación: supóngase que existe $\bar{x} \neq \bar{o}$ en la intersección del enunciado; se puede poner $\bar{x} = x_1 + ... + x_{i-1} + x_{i+1} + ... + x_p$. Toda suma $\Sigma \bar{u}_i$ se puede entonces poner también en la forma:

$$(\bar{u}_1 + \bar{x}_1) + ... + (\bar{u}_{i-1} + \bar{x}_{i-1}) + (\bar{u}_i - \bar{x}) + (\bar{u}_{i+1} + \bar{x}_{i+1}) + ... + (\bar{u}_p + x_p)$$

II.17. $U_1 \cap U_2 = O$. Cualquier $(y_1, y_2, ..., y_n) \in \mathbb{R}^n$ se puede poner:

$$(y_1, y_2, ..., y_n) = (x_1, x_2, ..., x_n) + (t, t, ..., t)$$

con

$$t = (y_1 + y_2 + ... + y_n) : n \quad \text{y} \quad x_i = y_i - t$$

$(i = 1, 2, ..., n)$.

II.18. $U_1 \cap U_2 = O$. Para cualquier $\varphi \in V$ se puede poner:

$$\varphi(x) = (a + bx) + f(x)$$

donde:

$$a = \frac{1}{2}[\varphi(1) + \varphi(-1)] \quad , \quad b = \frac{1}{2}[\varphi(1) - \varphi(-1)]$$

$f(x) = \varphi(x) - (a + bx)$ es tal que $f(1) = f(-1) = 0$.

II.19. *a*) Es independiente: la relación $\Sigma \lambda_i f(x)^i \equiv 0$ implica $\Sigma \lambda_i f(x_p)^i = 0$ para los n puntos x_p del enunciado; este sistema sólo tiene la solución nula (su determinante es de Wandermonde).

b) $V = U_1 \oplus U_2$ pues $U_1 \cap U_2 = O$ y todo $\varphi(x) \in V$ se puede poner $\varphi(x) = f(x) + h$ con

$$h = \int_0^1 \varphi(x)\,dx \quad \text{y} \quad f(x) = \varphi(x) - h$$

es tal que

$$\int_0^1 f(x)\,dx = 0$$

II.20. Recurrir a que una suma es directa si y sólo si $\bar{o} + \bar{o} + \cdots + \bar{o}$ es la única manera de descomponer $\bar{o}$ en suma de vectores de los subespacios sumandos.

II.21. $AB = BA \Rightarrow (AB)^t = B^t A^t = BA = AB \Rightarrow$
$\Rightarrow AB$ simétrica

AB simétrica $\Rightarrow AB = (AB)^t \Rightarrow AB = B^t A^t = BA$

II.22. Sean $f: \mathbb{R}^p \to \mathbb{R}^n$ y $g: \mathbb{R}^n \to \mathbb{R}^p$ aplicaciones lineales asociadas a A y B.

$$\text{rang}\,(AB) = \dim f(g(\mathbb{R}^n)) \quad , \quad \text{rang}\,A = \dim f(\mathbb{R}^p)$$
$$\text{rang}\,B = \dim g(\mathbb{R}^n)$$

$g(\mathbb{R}^n) \subset \mathbb{R}^p \Rightarrow$ 1.ª ; $f(g(\mathbb{R}^n)) \subset g(\mathbb{R}^n) \Rightarrow$ 2.ª

II.23. $A + N = A(I + A^{-1}N)$; $(A + N)^{-1} = (I + A^{-1}N)A^{-1}$; tómese $M = A^{-1}N$; $(A + N)^{-1} = A^{-1} - A^{-1}NA^{-1}$

II.24. Si; $I \in \mathcal{M}$; $A, B \in \mathcal{M} \Rightarrow AB^{-1} \in \mathcal{M}$

$$A^{-1} = \begin{bmatrix} 1 & -a & ac - b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix}$$

II.25. *a*) $y_1 = x_1 + 2x_2 + 3x_3 + x_4$, $y_2 = x_1 + x_2 + x_3$, $y_3 = -4x_1 - 2y_2 + 2x_4$

b) $(1, -2, 1, 0)$ y $(1, -1, 0, 1)$ forman base de Nuc (f)
$(1, 0, 2)$ y $(0, 1, -6)$ forman base de Im (f).

II.26. Si existiese $\bar{a} \neq \bar{o}$ en Nuc (f) y en Im (f), tómese una base de Im (f) que incluya a $\bar{a}$ y hállese una base de Im $(f \circ f)$.

II.27. *a*) $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

b) $\bar{e}_1 = (1, 0, 0)$, $\bar{e}_2 = (1, -1/2, 0)$, $\bar{e}_3 = (5, -4, 1)$
$\bar{u}_1 = (1, 3)$, $\bar{u}_2 = (0, 1)$

c) $g \circ f$ no puede ser la identidad y $f \circ g$ sí lo puede ser. Tómese para $g: \mathbb{R}^2 \to \mathbb{R}^3$, por ejemplo, la aplicación que en las bases $(\bar{u}_i)$ y $(\bar{e}_j)$ tiene por matriz a la traspuesta de D.

II.28. *a*) $B' = (\bar{e}_3, \bar{e}_2, -\bar{e}_1)$.
b) $B'' = (-\bar{e}_3, \bar{e}_2, \bar{e}_1)$.

II.29. *a*) V es subespacio de $\mathcal{M}_{nxn}$; S es forma lineal ya que $S(\lambda A + \mu B) = \lambda S(A) + \mu S(B)$. *b*) $\alpha = S(A)$. *c*) El rango de A no se altera al sumar a una de sus filas todas las demás; en $AA^{-1} = I$ y $A^{-1}A = I$ multiplicar por J y recurrir a la propiedad anterior; $S = (A^{-1}) = 1 : S(A)$.

II.30. 1. $\begin{bmatrix} \alpha & -h\beta \\ \beta & \alpha \end{bmatrix}$

2. Si $h > 0$, no hay; si $h \leq 0$, los divisores de cero se obtienen para $\alpha = \pm \beta\sqrt{-h}$.
3. Es cuerpo si $h \leq 0$.

II.31. 1. $A(x_1, x_2, x_3, \ldots, x_n) = x_1 I + x_2 E + x_3 E^2 + \cdots + x_n E^{n-1}$.
2. $\mathcal{M}$ tiene dimensión n; una de sus bases está formada por $A(1, 0, \ldots, 0)$, $A(0, 1, \ldots, 0)$, $\ldots$, $A(0, 0, \ldots, 1)$.
3. $A(1, -2, 1, 0, \ldots, 0)$.

II.32. 1. $\begin{bmatrix} 1 & -1 & 0 & -1 \\ 2 & -1 & -1 & -1 \\ 1 & -1 & 0 & -1 \\ -1 & 0 & 1 & 0 \end{bmatrix}$

2. $\begin{bmatrix} 0 & 0 & 1 & -1 \\ 2 & -1 & -1 & -2 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 1 \end{bmatrix}$

II.33. 1. $y_1 = x_1 + x_3 + x_4$
$y_2 = 2x_2 + 2x_3 + 4x_4$
$y_3 = x_1 - x_2 - x_4$
$y_4 = x_2 + x_3 + 2x_4$
$y_5 = -x_1 + 2x_2 + x_3 + 3x_4$

2. $y_1 = x'_1 - x'_2 + 3x'_3$
$y_2 = 2x'_1 + 6x'_3 - 2x'_4$
$y_3 = -x'_2 + x'_4$
$y_4 = x'_1 + 3x'_3 - x'_4$.
$y_5 = x'_1 + x'_2 + 3x'_3 - 2x'_4$

3. $(-1, -1, 1, 0)$ y $(-1, -2, 0, 1)$ forman base de Nuc (f)
$(1, 0, 1, 0, -1)$ y $(1, 2, 0, 1, 1)$ forman base de Im (f)
4. $V = \{(\alpha, \beta, 0, 0) \in \mathbb{R}^4 / \alpha, \beta \in \mathbb{R}\}$

II.34. 1. $\begin{bmatrix} 1 & 0 & 1 \\ -3 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

3. En V la nueva base $\left(\frac{3}{7} + \frac{1}{7}x + \frac{4}{7}x^2, -\frac{2}{3} - \frac{3}{7}x + \frac{2}{7}x^2, \frac{1}{7} + \frac{5}{7}x - \frac{1}{7}x^2\right)$

II.35. 1. $a = 1$

2. $P = \begin{bmatrix} 1 & 2 & -2 & -6 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 4 \end{bmatrix}$

$\overline{Q} = \begin{bmatrix} 1/2 & 0 & 0 \\ 2 & -2 & 1/2 \\ 1 & -1 & 0 \end{bmatrix}$

3. B: $(1, 0, 0, 0)$, $(2, 0, 0, -1)$, $(-2, 1/2, 0, 0)$, $(-6, 0, 1, 4)$
C: $(1/2, 2, 1)$, $(0, -2, -1)$, $(0, 1/2, 0)$

II.36. 1. $\alpha = 2$ o $\beta = 0$
2. Si $\alpha \neq 2$ y $\beta \neq 0$, Im $(f) = V$ y Nuc $(f) = O$
Si $\alpha = 2$ y $\beta \neq 0$,
Im $(f) = \mathcal{V}(2x + \beta x^2, \beta + x)$,
Nuc $(f) = \mathcal{V}(1, 1, -1)$
Si $\beta = 0$, Im $(f) = \mathcal{V}(x)$,
Nuc $(f) = \mathcal{V}((1, 2, 0), (0, \alpha - 1, 1))$
3. Si $\alpha \neq 2$, $f(S) = \mathcal{V}(x)$; si $\alpha = 2$, $f(S) = O$.

II.37. 1. $\begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Nuc $(h) = \mathcal{V}(1, -1, 0)$
Im $(h) = \mathcal{V}((0, 1, 1), (1, 0, 0))$

2. $\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ 2 & 2 & 1 & -2 \\ 1 & 1 & 1 & -1 \end{bmatrix}$

Nuc $(k) = \mathcal{V}((1, -1, 0, 0), (1, 0, 0, 1))$
Im $(k) = \mathcal{V}((1, 1, 2, 1), (1, 0, 1, 1))$

3. Im (f) = Im (k), Nuc (f) = Nuc (h)

II.38. Nuc $(\varphi) = O$, Im (φ) = Im (f)

II.39. 1. Nuc $(F) = \{(a, a, -a) \in \mathbb{R}^3 / a \in \mathbb{R}\}$,
Im $(F) = \{\alpha + \beta \operatorname{sen}^2 x / \alpha, \beta \in \mathbb{R}\}$
Nuc $(G) = \{f \in V / f(0) = f(-\pi/2) = 0\}$,
Im $(G) = \mathbb{R}$
2. $F^{-1}(U) = \{(a, a, c) \in \mathbb{R}^3 / a, c \in \mathbb{R}\}$
3. En $\mathbb{R}^3$: $(1, 0, 0)$, $(1, 1, 0)$ y $(-1, -1, 1)$.
En $F(\mathbb{R}^3)$: $\operatorname{sen}^2 x$, 1

4. $(G \circ F)(a, b, c) = (b + c, -a + c)$,
$\text{Nuc}\,(G \circ F) = \{(a, -a, a) \in \mathbb{R}^3 / a \in \mathbb{R}\}$

II.40. Nótese que $f \in \mathcal{V} \Leftrightarrow \text{Im}\,(f) \subset \text{Nuc}\,(\varphi)$. El espacio $\mathcal{V}$ es isomorfo al $\mathcal{L}(V, \text{Nuc}\,(\varphi))$, luego tiene dimensión $n \cdot d$, donde d es la dimensión de Nuc (φ).

II.41. Tómense bases en las que la matriz de f es la canónica (de equivalencia) C_r. Póngase $C_r = 2I + (C_r - 2I)$, donde I es la unidad $n \times n$; $2I$ y $C_r - 2I$ son regulares.

II.42. Tómense en V y W bases en las que la matriz de f es la canónica (de equivalencia) C_r; f inyectiva si y sólo si $r = n$. Si $r = n$, g puede ser el homomorfismo de matriz C_r^t. Si $g \circ f = i$, la matriz de f no puede tener una columna nula, luego $r = n$.

II.43. Como en el ejercicio anterior, recurramos a C_r; f es sobreyectiva si y sólo si $r = n$. Si $r = m$, g puede ser el homomorfismo de matriz C_r^t. Si $f \circ g = i$, la matriz de f no puede tener una fila nula, luego $r = m$.

II.44. Buscando f^{-1} de la forma $\alpha i + \beta f$, se obtiene $f^{-1} = -a - bf$.

II.45. Recuérdese que dim Nuc (f) + dim Im $(f) = n$ y que $f \circ f = o$ equivale a Im $(f) \subset$ Nuc (f).

II.46. *a*) $\varphi(x, y, z) = 6x + y - 2z$
b) Coordenadas de $\varphi = (6, 1, -2)$

II.47. 1. $\varphi(x_1, x_2, x_3) = x_1 + ax_2$
$\psi(x_1, x_2, x_3) = 3x_2 + x_3$
$\eta(x_1, x_2, x_3) = 2x_1 + 2x_2 + 2bx_3$
2. $\varphi : (1, a, 0)$, $\psi : (0, 3, 1)$, $\eta : (2, 2, 2b)$
3. $a + 3b = 1$

II.48. $\varphi_i(\bar{u}_j) = 1$, si $i = j$; $\varphi_i(\bar{u}_j) = 0$, si $i \neq j$

$$\varphi_1(x, y, z) = x \quad , \quad \varphi_2(x, y, z) = 7x - 2y - 3z$$
$$\varphi_3(x, y, z) = -2x + y + z$$

II.49. $p_1(x) = 2 - 2x$, $p_2(x) = -(1/2) + t$.

II.50. Como $f \neq o$, existe $\bar{v}_0 \in V$ tal que $f(\bar{v}_0) \neq 0$; se verifica que $\mathcal{V}(\bar{v}_0) \oplus U = V$.

II.51. Se sabe que existe $\bar{v}_0 \in V$ tal que $\mathcal{V}(\bar{v}_0) \oplus U = V$; para cada $\bar{x} \in V$, existen unos únicos $\lambda \in K$ y $\bar{u} \in V$ tales que $\bar{x} = \lambda \bar{v}_0 + \bar{u}$. Tómese $f(\bar{x}) = \lambda$.

PARTE III

AQUÍ, cuando las variables (que fueron llamadas independientes) varían, que ese es su cometido, y arrastran en su variar a las funciones (que antaño también se llamaban «las variables dependientes»), va a acontecer que la variación de éstas, las funciones, será cuadrática (respecto del variar de la variable, que será nuestra referencia sobre el modo de variar).
Hay que hacer notar que, aun pareciendo paradójico, no por ello nos separamos de lo que es la linealidad, que nos acompañará hasta el final. Esto de estudiar lo cuadrático no es capricho. Es necesidad; sin ello, nada podríamos decir sobre las cosas del «medir», y ya es hora de encararnos con las distancias y con los ángulos.
Recuerde el lector cuando recitaba el teorema de Pitágoras y decía: «el "cuadrado" de la hipotenusa es igual a la suma de los "cuadrados" de los catetos».

7. Formas cuadráticas.
8. Espacios vectoriales euclídeos. Ejercicios y problemas.

CAPITULO

7

Formas cuadráticas

Hay dos maneras de abordar las cuestiones que componen esta tercera parte del libro. Se podría empezar con el estudio de los espacios vectoriales euclídeos, esto es, introduciendo el concepto de producto escalar, con el que un espacio vectorial real pasa a ser euclídeo, y después generalizar, hablando de las formas cuadráticas. A este camino se le suponen ventajas de tipo didáctico, pues va de lo particular a lo general, pero obliga a reiterar no pocas cuestiones que, con distinto alcance y quizá cambiando de nombre, es necesario estudiar en uno y otro caso.

Acontece, además, que el paso de un producto escalar a una forma cuadrática general se produce relajando las exigencias de aquel, esto es, prescindiendo de algunos de sus axiomas. Con ello, aquél que estudió primero el producto escalar y se acostumbró a su modo de funcionar, propenderá a creer que las formas cuadráticas se comportan, en todo, del modo que él ya conoce y, por esto, ha de costarle mucho discernir cuáles, de entre todas las propiedades que le son familiares, siguen siendo válidas y cuáles ya no lo son.

Consecuentes con lo anterior, hablaremos en primer lugar sobre las formas cuadráticas, en este capítulo, y dejaremos para después el estudio del producto escalar, que trataremos en el próximo.

α FORMAS CUADRATICAS. CONJUGACION

7.1. FORMAS BILINEALES

Antes de abordar el estudio de las formas cuadráticas, es necesario decir algunas cosas sobre las formas bilineales, que son soporte de aquellas.

Se llaman aplicaciones bilineales aquellas funciones de dos variables que son lineales respecto de cada una de ellas, esto es: para que una aplicación

$$(\bar{x}, \bar{y}) \longmapsto f(\bar{x}, \bar{y})$$

sea bilineal se requiere que, al tomar $\bar{x} = \bar{a}$ o $\bar{y} = \bar{b}$, valores fijos arbitrariamente elegidos de las variables, las funciones

$$\bar{x} \longmapsto f(\bar{x}, \bar{b}) \quad \text{e} \quad \bar{y} \longmapsto f(\bar{a}, \bar{y})$$

sean lineales. Más exactamente, se conviene en decir:

FORMA BILINEAL SOBRE UN ESPACIO VECTORIAL

[112]

Sea V un espacio vectorial sobre el cuerpo K (*) y considérese una aplicación f, que a cada par de vectores le asigna un escalar (**), es decir:

$$f: V \times V \to K, \quad (\bar{x}, \bar{y}) \mapsto f(\bar{x}, \bar{y})$$

Se dice que f es una forma bilineal sobre V si se verifica que:

$$f(\lambda\bar{x} + \lambda'\bar{x}', \bar{y}) = \lambda f(\bar{x}, \bar{y}) + \lambda' f(\bar{x}', \bar{y}), \qquad \forall\, \bar{x}, \bar{x}', \bar{y} \in V, \forall\, \lambda, \lambda' \in K$$

$$f(\bar{x}, \mu\bar{y} + \mu'\bar{y}') = \mu f(\bar{x}, \bar{y}) + \mu' f(\bar{x}, \bar{y}'), \qquad \forall\, \bar{x}, \bar{y}, \bar{y}' \in V, \forall\, \mu, \mu' \in K$$

El conjunto de las formas bilineales sobre V, que denotaremos por $\mathcal{B}(V)$, es un espacio vectorial, sobre K, respecto de las operaciones usuales (suma de funciones y producto de función por escalar).

(*) Para nosotros, los casos importantes son $K = \mathbb{R}$ y $K = \mathbb{C}$.

(**) A la función f se la llama «forma» debido a que toma sus valores en el cuerpo K de los escalares.

DEMOSTRACION Y COMENTARIOS

1. Según ya dijimos, estas aplicaciones se llaman bilineales porque son lineales respecto de cada una sus variables; es decir, debido a que las dos condiciones que se las exige significan que son lineales las siguientes aplicaciones:

   Para cada $\bar{y} \in V$ fijo, la aplicación $f(_, \bar{y}): V \to K$, $\bar{x} \mapsto f(\bar{x}, \bar{y})$.
   Para cada $\bar{x} \in V$ fijo, la aplicación $f(\bar{x}, _): V \to K$, $\bar{y} \mapsto f(\bar{x}, \bar{y})$.

2. Las dos condiciones de la definición pueden resumirse en la siguiente: para cualesquiera $\bar{x}, \bar{x}', \bar{y}, \bar{y}' \in V$ y $\lambda, \lambda', \mu, \mu' \in K$ es:

   $$f(\lambda\bar{x} + \lambda'\bar{x}', \mu\bar{y} + \mu'\bar{y}') = \lambda\mu f(\bar{x}, \bar{y}) + \lambda\mu' f(\bar{x}, \bar{y}') + \lambda'\mu f(\bar{x}', \bar{y}) + \lambda'\mu' f(\bar{x}', \bar{y}')$$

3. Si f es una forma bilineal, es evidente que, para cualesquiera vectores $\bar{x}, \bar{y} \in V$, se verifica:

   $$f(\bar{x}, \bar{o}) = f(\bar{o}, \bar{y}) = 0$$
   $$f(-\bar{x}, \bar{y}) = f(\bar{x}, -\bar{y}) = -f(\bar{x}, \bar{y})$$

4. Para verificar que $\mathcal{B}(V)$ es un espacio vectorial, hay que probar que si $f, g \in \mathcal{B}(V)$ y $h, k \in K$, entonces $hf + kg \in \mathcal{B}(V)$. Así ocurre, ya que para cualesquiera $\bar{x}, \bar{x}', \bar{y}, \bar{y}' \in V$ y $\lambda, \lambda', \mu, \mu' \in V$ es:

- $(hf+kg)(\lambda\bar{x}+\lambda'\bar{x}', \bar{y}) = hf(\lambda\bar{x}+\lambda'\bar{x}', \bar{y}) + kg(\lambda\bar{x}+\lambda'\bar{x}', \bar{y}) =$
  $= h[\lambda f(\bar{x}, \bar{y}) + \lambda' f(\bar{x}', \bar{y})] + k[\lambda g(\bar{x}, \bar{y}) + \lambda' g(\bar{x}', \bar{y})] =$
  $= \lambda(hf+kg)(\bar{x}, \bar{y}) + \lambda'(hf+kg)(\bar{x}', \bar{y})$
- $(hf+kg)(x, \mu\bar{y}+\mu'\bar{y}') = \mu(hf+kg)(\bar{x}, \bar{y}) + \mu'(hf+kg)(\bar{x}, \bar{y}')$
  (esta segunda se comprueba de igual manera que la primera).

EJEMPLOS

1. La aplicación $f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, que a cada par de vectores $\bar{x} = (x_1, x_2)$, $\bar{y} = (y_1, y_2)$ le asigna el valor

$$f(\bar{x}, \bar{y}) = f[(x_1, x_2), (y_1, y_2)] = 3x_1y_1 - 6x_1y_2 + 5x_2y_1 + 4x_2y_2$$

es una forma bilineal sobre $\mathbb{R}^2$. Nótese que esta forma bilineal se puede expresar matricialmente poniendo:

$$f(\bar{x}, \bar{y}) = [x_1\ x_2]\begin{bmatrix} 3 & -6 \\ 5 & 4 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

2. Si $\varphi_1: V \to K$ y $\varphi_2: V \to K$ son dos formas lineales, entonces la aplicación

$$f: V \times V \to K, \quad (\bar{x}, \bar{y}) \mapsto f(\bar{x}, \bar{y}) = \varphi_1(\bar{x})\varphi_2(\bar{y})$$

es una forma bilineal.

3. Llamando V al espacio vectorial de las funciones de $[0, 1]$ en $\mathbb{R}$ que tienen derivada continua, la aplicación

$$V \times V \to \mathbb{R}, \quad (f, g) \mapsto \int_0^1 f(x)g'(x)\,dx$$

es una forma bilineal, lo que se comprueba fácilmente recurriendo a la linealidad de la integral y de la derivada.

OBSERVACION

Aun cuando aquí no hagamos uso de ella, tiene interés conocer la siguiente generalización, en la que «lo bilineal» entre espacios vectoriales se considera con la mayor generalidad:

Dados tres espacios vectoriales U, V y W, todos ellos sobre el mismo cuerpo K, se dice que una aplicación $f: U \times V \to W$ es bilineal si, para cualesquiera $\bar{u}, \bar{u}' \in U$, $\bar{v}, \bar{v}' \in V$ y $\lambda, \lambda' \in K$, se verifica que:

$$f(\lambda\bar{u}+\lambda'\bar{u}', \bar{v}) = \lambda f(\bar{u}, \bar{v}) + \lambda' f(\bar{u}', \bar{v})$$
$$f(\bar{u}, \lambda\bar{v}+\lambda'\bar{v}') = \lambda f(\bar{u}, \bar{v}) + \lambda' f(\bar{u}, \bar{v}')$$

EJERCICIO

Sea $f: V \times V \to K$ una forma bilineal. Pruébese que la aplicación $f^*: V \times V \to K$ definida mediante

$$f^*(\bar{x}, \bar{y}) = f(\bar{y}, \bar{x}), \quad \forall \bar{x}, \bar{y} \in V$$

es una forma bilineal (a f^* se la llama adjunta o traspuesta de f).

RESOLUCION

Para cualesquiera $\bar{x}, \bar{x}', \bar{y}, \bar{y}' \in V$ y $\lambda, \lambda' \in K$ es:

$$f^*(\lambda\bar{x} + \lambda'\bar{x}', \bar{y}) = f(\bar{y}, \lambda\bar{x} + \lambda'\bar{x}') = \lambda f(\bar{y}, \bar{x}) + \lambda' f(\bar{y}, \bar{x}') = \\ = \lambda f^*(\bar{x}, \bar{y}) + \lambda' f^*(\bar{x}', \bar{y})$$

$$f^*(\bar{x}, \lambda\bar{y} + \lambda'\bar{y}') = f(\lambda\bar{y} + \lambda'\bar{y}', \bar{x}) = \lambda f(\bar{y}, \bar{x}) + \lambda' f(\bar{y}', \bar{x}) = \\ = \lambda f^*(\bar{x}, \bar{y}) + \lambda' f^*(\bar{x}, \bar{y}')$$

EXPRESION ANALITICA DE UNA FORMA BILINEAL

[113]

Sea V un espacio vectorial, sobre el cuerpo K y sea $f: V \times V \to K$ una forma bilineal. Si V tiene dimensión n y si $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una base de V, entonces para cualesquiera $\bar{x}, \bar{y} \in V$ es:

$$f(\bar{x}, \bar{y}) = \sum_{i,j=1}^{n} a_{ij} x_i y_j \overset{(*)}{=} X'AY = [x_1 x_2 ... x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

donde las a_{ij} son los escalares $a_{ij} = f(\bar{e}_i, \bar{e}_j)$; donde x_i e y_j denotan a las coordenadas de $\bar{x}$ e $\bar{y}$, esto es, $\bar{x} = \Sigma x_i \bar{e}_i$ e $\bar{y} = \Sigma y_j \bar{e}_j$; donde $A = [a_{ij}]$ es matriz $n \times n$; y donde X e Y son las matrices columna de las coordenadas de $\bar{x}$ e $\bar{y}$.

Se dice que A es la *matriz (de Gram) de la forma bilineal* f en la base B. La correspondencia $f \mapsto A$ es un isomorfismo del espacio vectorial $\mathfrak{B}(V)$, de las formas bilineales sobre el espacio V de dimensión n, en el espacio vectorial $\mathcal{M}_n$, de las matrices cuadradas $n \times n$.

(*) Nótese que $X'AY$ es una matriz de tamaño 1×1 cuyo único elemento es el escalar $\Sigma a_{ij} x_i y_j$, con el que se identifica dicha matriz.

DEMOSTRACION

1. De acuerdo con la definición de forma bilineal, se verifica que

$$f(\bar{x}, \bar{y}) = f\left(\sum_i x_i \bar{e}_i, \bar{y}\right) = \sum_i x_i f(\bar{e}_i, \bar{y}) = \sum_i x_i f\left(\bar{e}_i, \sum_j y_j \bar{e}_j\right) =$$

$$= \sum_i x_i \left(\sum_j y_j f(\bar{e}_i, \bar{e}_j)\right) = \sum_i \sum_j x_i y_j a_{ij} = \sum_{i,j} a_{ij} x_i y_j$$

2. En el supuesto de disponer de la base B (fija), la correspondencia que a cada forma bilineal $f \in \mathcal{B}(V)$ le asigna su matriz $A \in \mathcal{M}_n$ es biyectiva. Nótese que, dada A, la aplicación $(\bar{x}, \bar{y}) \rightarrowtail X^t A Y$ es una forma bilineal, como se comprueba fácilmente.

   Esta biyección $f \rightarrowtail A$ es un isomorfismo entre los espacios vectoriales $\mathcal{B}(V)$ y $\mathcal{M}_n$, ya que para cualesquiera $f, g \in \mathcal{B}(V)$ y $\lambda, \mu \in K$ es:

$$\text{matriz de } [\lambda f + \mu g] = \lambda[\text{matriz de } f] + \mu[\text{matriz de } g]$$

   en efecto: si A y B son las matrices de f y g, la matriz de $\lambda f + \mu g$ es la $\lambda A + \mu B$, ya que (para cualesquiera $\bar{x}, \bar{y} \in V$) es:

$$(\lambda f + \mu g)(\bar{x}, \bar{y}) = \lambda f(\bar{x}, \bar{y}) + \mu g(\bar{x}, \bar{y}) = \lambda X^t A Y + \mu X^t B Y = X^t(\lambda A + \mu B)Y$$

EJEMPLO

Sea $f: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ la forma bilineal definida por

$$f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 2x_1y_1 - 3x_1y_2 + x_2y_1 + 6x_1y_3 + 4x_3y_2 - x_3y_3$$

La expresión matricial de f en la base canónica de $\mathbb{R}^3$ es:

$$f((x_1, x_2, x_3), (y_1, y_2, y_3)) = [x_1 x_2 x_3] \begin{bmatrix} 2 & -3 & 6 \\ 1 & 0 & 0 \\ 0 & 4 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = X^t A Y$$

La matriz A' de f en la nueva base $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$, donde $\bar{u}_1 = (2, 0, 0)$, $\bar{u}_2 = (1, 2, 0)$ y $\bar{u}_3 = (-3, 1, 1)$, será:

$$A' = \begin{bmatrix} f(\bar{u}_1, \bar{u}_1) & f(\bar{u}_1, \bar{u}_2) & f(\bar{u}_1, \bar{u}_3) \\ f(\bar{u}_2, \bar{u}_1) & f(\bar{u}_2, \bar{u}_2) & f(\bar{u}_2, \bar{u}_3) \\ f(\bar{u}_3, \bar{u}_1) & f(\bar{u}_3, \bar{u}_2) & f(\bar{u}_3, \bar{u}_3) \end{bmatrix} = \begin{bmatrix} 8 & -8 & -6 \\ 8 & -2 & -9 \\ -10 & 21 & 9 \end{bmatrix}$$

OBSERVACIONES

1. Una forma bilineal $f \in \mathcal{B}(V)$, donde el espacio V tiene dimensión n, queda determinada en cuanto se conozcan los n^2 escalares $a_{ij} = f(\bar{e}_i, \bar{e}_j)$, siendo

$B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ una base de V. Dicho de otro modo, la matriz $A = [a_{ij}]$ caracteriza a la forma bilineal f.

2. Conviene notar que para expresar matricialmente el valor de $f(\bar{x}, \bar{y})$, además de recurrir a la matriz A, como se ha hecho y haremos en el futuro, también se puede hacer uso de la matriz A^t; para ello, como X^tAY es una matriz 1×1 y, por tanto, coincide con su traspuesta, se puede poner

$$f(\bar{x}, \bar{y}) = X^tAY = (X^tAY)^t = Y^tA^tX$$

Hay quienes prefieren poner Y^tA^tX, en lugar de X^tAY, y dicen que la matriz de f es A^t, en lugar de A.

3. Como, si dim $V = n$, el espacio vectorial $\mathcal{B}(V)$ es isomorfo al $\mathcal{M}_n$ y este último tiene dimensión n^2, resulta que $\mathcal{B}(V)$ tiene dimensión n^2. La base usual de $\mathcal{B}(V)$ está formada por las n^2 formas bilineales f_{ij} (para $i, j = 1, 2, ..., n$) definidas mediante:

$$f_{ij}(\bar{e}_h, \bar{e}_k) = \begin{cases} 1 & \text{si } h = i \text{ y } k = j \\ 0 & \text{si } h \neq i \text{ o } k \neq j \end{cases}$$

CAMBIO DE BASE. MATRICES CONGRUENTES

La matriz A asociada a una forma bilineal $f \in \mathcal{B}(V)$ depende de la base que se tome en V. Interesa analizar cómo varía la matriz A con la base; las matrices asociadas a una misma forma bilineal f, respecto de las distintas bases de V, se dice que son congruentes entre sí:

[114]

Sea $f: V \times V \to K$ una forma bilineal, donde V es un espacio vectorial, sobre el cuerpo K, que tiene dimensión n. Si A es la matriz asociada a f en una primera base de V y si A' es la matriz asociada a f respecto de otra base, entonces se verifica que:

$$A' = P^tAP$$

donde P es la matriz del correspondiente cambio de coordenadas(*).

Dos matrices, A y A', cuadradas y del mismo tamaño $n \times n$, se dicen *matrices congruentes* si existe alguna matriz regular P tal que $A' = P^tAP$. Dos matrices son congruentes si están asociadas, ambas, a una misma forma bilineal, respecto de distintas bases. Para el conjunto $\mathcal{M}_n$, de las matrices $n \times n$, la congruencia de matrices es una relación de equivalencia.

(*) Esta matriz P es la que permite poner $X = PX'$, donde X y X' son las matrices columna de coordenadas, de un mismo vector genérico, en la primera y la segunda bases, respectivamente.

DEMOSTRACION

1. Sean $\bar{x}, \bar{y} \in V$ dos vectores cualesquiera; llamemos X y X' e Y e Y' a las matrices columna de coordenadas de $\bar{x}$ e $\bar{y}$ en la primera y en la segunda base, respectivamente, de manera que $X = PX'$ e $Y = PY'$. Podemos, entonces poner:

$$f(\bar{x},\bar{y}) = X^tAY = (PX')^tA(PY') = (X'^tP^t)A(PY') = X'^t(P^tAP)Y'$$

Según este resultado, la matriz A' asociada a f en la nueva base es:

$$A' = P^tAP \quad (P \text{ regular})$$

2. El recíproco es evidente: si $A' = P^tAP$, donde $P = [p_{ij}]$ es una matriz regular, entonces A y A' están asociadas a una misma forma bilineal. En efecto: sea $f \in \mathcal{B}(K^n)$ la forma bilineal que en la base canónica $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de K^n tiene asociada a la matriz A; en la base $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$, donde $\bar{u}_j = \Sigma p_{ij}\bar{e}_i$ (se suma para $i = 1, 2, ..., n$), la nueva matriz de f es la P^tAP.
3. Para comprobar que la congruencia de matrices es una relación de equivalencia en $\mathcal{M}_n$, comprobemos que es reflexiva, simétrica y transitiva:

   - Es reflexiva, pues toda matriz A es congruente consigo misma debido a que $A = I^tAI$, donde I es la matriz unidad, que es regular.
   - Es simétrica, pues si A es congruente con A', entonces A' es congruente con A, ya que:

$$A' = P^tAP \ (P \text{ regular}) \quad \Rightarrow \quad A = (P^{-1})^tA'(P^{-1})$$

   donde P^{-1} es regular.
   - Es transitiva, pues si A es congruente con A' y A' lo es con A'', entonces A es congruente con A'', ya que:

$$\left.\begin{aligned} A' &= P^tAP \ (P \text{ regular}) \\ A'' &= Q^tA'Q \ (Q \text{ regular}) \end{aligned}\right\} \quad \Rightarrow \quad A'' = (PQ)^tA(PQ)$$

   donde PQ es regular.

EJEMPLO

Sea $f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ la forma bilineal:

$$f(\bar{x}, \bar{y}) = f((x_1, x_2), (y_1, y_2)) = 8x_1y_1 + 9x_1y_2 + 10x_2y_1 + 11x_2y_2$$

Al expresar esta forma bilineal en lo base $(\bar{u}_1, \bar{u}_2)$, donde

$$\bar{u}_1 = (-1, 1) \quad \text{y} \quad \bar{u}_2 = (4, -3),$$

la matriz de f será:

$$P^tAP = \begin{bmatrix} -1 & 1 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} 8 & 9 \\ 10 & 11 \end{bmatrix}\begin{bmatrix} -1 & 4 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix}$$

Es decir, la expresión de f en las nuevas coordenadas $\bar{x}(x'_1, x'_2)$ e $\bar{y}(y'_1, y'_2)$ es:

$$f(\bar{x}, \bar{y}) = 2x'_1y'_2 + x'_2y'_1 - x'_2y'_2$$

FORMAS BILINEALES SIMETRICAS Y ANTISIMETRICAS

[115]

Sea V un espacio vectorial, sobre un cuerpo K, y $f: V \times V \to K$ una forma bilineal. Se dice que:

- f es *simétrica* si $f(\bar{y}, \bar{x}) = f(\bar{x}, \bar{y})$, $\forall\, \bar{x}, \bar{y} \in V$.
- f es *antisimétrica* si $f(\bar{y}, \bar{x}) = -f(\bar{x}, \bar{y})$, $\forall \bar{x}, \bar{y} \in V$

Toda forma bilineal es(*) suma de una forma bilineal simétrica f_s y otra antisimétrica f_a:

$$f = f_s + f_a$$

donde

$$f_s(\bar{x}, \bar{y}) = \tfrac{1}{2}\,[f(\bar{x}, \bar{y}) + f(\bar{y}, \bar{x})] \quad \text{y} \quad f_a(\bar{x}, \bar{y}) = \tfrac{1}{2}\,[f(\bar{x}, \bar{y}) - f(\bar{y}, \bar{x})]$$

Los conjuntos $\mathcal{B}_s(V)$ y $\mathcal{B}_a(V)$, de las formas bilineales simétricas y antisimétricas, son(*) subespacios suplementarios del espacio $\mathcal{B}(V)$, de las formas bilineales.

Si el espacio V tiene dimensión finita y si A es la matriz de una forma bilineal $f: V \times V \to K$, respecto de una base cualquiera, entonces:

[f simétrica $\Leftrightarrow$ A simétrica] y [f antisimétrica $\Leftrightarrow$ A antisimétrica]

Si A es la matriz de una forma bilineal f, las matrices A_s y A_a de f_a y f_s son (*):

$$A_s = \tfrac{1}{2}\,[A + A^t] \quad \text{y} \quad A_a = \tfrac{1}{2}\,[A - A^t]$$

(nótese que $A = A_s + A_a$).

(*) Aquí se requiere que el cuerpo K tenga «característica distinta de 2», lo que significa que $1 + 1 \neq 0$ (donde 1 es la unidad de K y 0 su elemento nulo). Esto es cierto, obviamente, para $K = \mathbb{R}$ y $K = \mathbb{C}$, que son los casos de mayor interés para nosotros.

DEMOSTRACION

1. Nótese que f_s y f_a son formas bilineales, pues son suma y diferencia de formas bilineales y éstas forman espacio vectorial (véase [112]). Es evidente que f_s es simétrica y f_a antisimétrica; también es obvio que su suma es f.

   Hay en esto un posible fallo que, hasta ahora, hemos intentado ocultar. Se trata de lo siguiente: todo lo dicho funciona sin problemas si se pone:

$$f(\bar{x}, \bar{y}) + f(\bar{x}, y) = [f(\bar{x}, \bar{y}) + f(\bar{y}, \bar{x})] + [f(\bar{x}, \bar{y}) - f(\bar{y}, \bar{x})]$$

   Como el primer miembro es $(1 + 1)f(\bar{x}, \bar{y}) = 2f(\bar{x}, \bar{y})$, si $1 + 1 \neq 0$ (o sea, si $2 \neq 0$), todo cuanto se ha obtenido anteriormente es correcto, pues se obtiene de dividir por $2 \neq 0$ en la última igualdad. Pero si el cuerpo K tuviese característica 2, esto es, si fuese $1 + 1 = 0$, entonces los resultados anteriores no serían ciertos.

2. Para comprobar que $\mathcal{B}_s(V)$ y $\mathcal{B}_a(V)$ son subespacios, hemos de ver que las combinaciones lineales $\lambda f + \mu g$ de dos formas bilineales simétricas o antisimétricas f y g son, respectivamente, simétricas o antisimétricas; así ocurre, ya que para cualesquiera $\bar{x}, \bar{y} \in V$ es:

   - (f y g simétricas) $\Rightarrow (\lambda f + \mu g)(\bar{x}, \bar{y}) = \lambda f(\bar{x}, \bar{y}) + \mu g(\bar{x}, \bar{y}) =$
     $= \lambda f(\bar{y}, \bar{x}) + \mu g(\bar{y}, \bar{x}) = (\lambda f + \mu g)(\bar{y}, \bar{x})$.
   - (f y g antisimétricas) $\Rightarrow (\lambda f + \mu g)(\bar{x}, \bar{y}) = \lambda f(\bar{x}, \bar{y}) + \mu g(\bar{x}, \bar{y}) =$
     $= -\lambda f(\bar{y}, \bar{x}) - \mu g(\bar{y}, \bar{x}) = -(\lambda f + \mu g)(\bar{y}, \bar{x})$.

3. Los subespacios $\mathcal{B}_s(V)$ y $\mathcal{B}_a(V)$ son suplementarios en $\mathcal{B}(V)$ ya que:

   - Su suma es $\mathcal{B}(V)$, pues toda forma bilineal f se puede poner $f = f_s + f_a$ (como se vio en 1), con $f_s \in \mathcal{B}_s(V)$ y $f_a \in \mathcal{B}_a(V)$.
   - Su intersección es nula, pues si f fuese simétrica y antisimétrica, entonces para cualesquiera $\bar{x}, \bar{y} \in V$ sería $f(\bar{x}, \bar{y}) = f(\bar{y}, \bar{x}) = -f(\bar{y}, \bar{x})$, luego $2f(\bar{x}, \bar{y}) = 0$ de donde $f(\bar{x}, \bar{y}) = 0$ (se supone que $2 \neq 0$, o sea, que K tiene característica distinta de 2) para cualesquiera $\bar{x}, \bar{y} \in V$, es decir, $f = o$, como se había dicho.

4. En el caso de dimensión finita, las formas f simétricas y antisimétricas son las que tienen matriz $A = [a_{ij}]$ simétrica y antisimétrica, respectivamente. En efecto:

   - f simétrica $\Leftrightarrow f(\bar{x}, \bar{y}) = f(\bar{y}, \bar{x}), \forall \bar{x}, \bar{y} \Leftrightarrow \sum_{ij} a_{ij} x_i y_j = \sum_{ij} a_{ij} y_i x_j, \forall \bar{x}, \bar{y}$
     $\Leftrightarrow \sum_{ij} a_{ij} x_i y_j = \sum_{ij} a_{ji} y_j x_i, \forall \bar{x}, \bar{y} \Leftrightarrow \sum_{ij} (a_{ij} - a_{ji}) x_i y_j, \forall \bar{x}, \bar{y} \Leftrightarrow$
     $\Leftrightarrow a_{ij} = a_{ji}, \forall i, j \Leftrightarrow A = A^t \Leftrightarrow A$ simétrica.
   - f antisimétrica $\Leftrightarrow a_{ij} = -a_{ji}, \forall i, j \Leftrightarrow A = -A^t \Leftrightarrow A$ antisimétrica.

5. Si A es la matriz de f, respecto de cierta base, y llamando X e Y a las matrices columna de coordenadas de los vectores genéricos $\bar{x}, \bar{y} \in V$ se tiene (nótese que $Y^t AX$ es un matriz 1×1, luego coincide con su traspuesta):

$$f_s(\bar{x}, \bar{y}) = \tfrac{1}{2}\,[f(\bar{x}, \bar{y}) + f(\bar{y}, \bar{x})] = \tfrac{1}{2}\,[X^tAY - Y^tAX] = \tfrac{1}{2}\,[X^tAY - X^tA^tY] =$$

$$= X^t[\tfrac{1}{2}\,(A + A^t)]Y \quad \Rightarrow \quad \tfrac{1}{2}\,(A + A^t) \text{ es la matriz de } f_s$$

Análogamente se procede con f_a y se obtiene la matriz $\frac{1}{2}(A - A^t)$.

EJEMPLOS

1. Considérese la forma bilineal $f: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ dada por:

$$f(\bar{x}, \bar{y}) = f((x_1, x_2, x_3), (y_1, y_2, y_3)) = \alpha x_1y_1 + \beta x_1y_2 + 6x_2y_1 + \gamma x_2y_2 + \delta x_1y_3 -$$

$$- 2x_3y_1 + \varepsilon x_3y_3 - 4x_3y_2 + \zeta x_3y_3 =$$

$$= [x_1\,x_2\,x_3]\begin{bmatrix} \alpha & \beta & \delta \\ 6 & \gamma & \varepsilon \\ -2 & 4 & \zeta \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = X^tAY$$

La forma f será simétrica si se toma $\beta = 6$, $\delta = -2$ y $\varepsilon = 4$, pudiendo ser α, β y ζ cualesquiera. La forma f será antisimétrica si se toma $\alpha = 0$, $\gamma = 0$, $\zeta = 0$, $\beta = -6$, $\delta = 2$ y $\varepsilon = -4$. Al tomar para los parámetros valores cualesquiera, la descomposición de f en suma de una forma simétrica y otra antisimétrica se obtiene, recurriendo a sus matrices, poniendo $A = A_a + A_s$ con:

$$A_s = \tfrac{1}{2}\,(A + A^t) = \begin{bmatrix} \alpha & 3 + \beta/2 & -1 + \delta/2 \\ 3 + \beta/2 & \gamma & 2 + \varepsilon/2 \\ -1 + \delta/2 & 2 + \varepsilon/2 & \zeta \end{bmatrix}$$

$$A_a = \tfrac{1}{2}\,(A - A^t) = \begin{bmatrix} 0 & -3 + \beta/2 & 1 + \delta/2 \\ -3 + \beta/2 & 0 & -2 + \varepsilon/2 \\ 1 + \delta/2 & -2 + \varepsilon/2 & 0 \end{bmatrix}$$

2. Sea V el espacio vectorial de las funciones continuas de $[0, 1]$ en $\mathbb{R}$. La aplicación

$$V \times V \to \mathbb{R}, \quad (f, g) \mapsto \int_0^1 f(x)g(x)\,dx$$

es una forma bilineal simétrica.

7.2. FORMAS CUADRATICAS

La variación de tipo cuadrático no nos debe resultar extraña, pues es la que gobierna en los polinomios homogéneos de segundo grado, de cualquier número de variables, como el

$$\omega(x, y) = ax^2 + bxy + cy^2 \quad \text{(de dos variables)}$$

Para estudiar tales funciones, las cuadráticas, se recurrirá a las formas bilineales simétricas, de las que aquellas se derivan. Así, para la anterior función $\omega(x, y)$, echaremos mano de la aplicación bilineal

$$f((x, y), (x', y')) = axx' + \frac{b}{2}(xy' + x'y) + cyy'$$

Nótese que esta f se ha elegido de manera que al tomar $x = x'$ e $y = y'$, se convierte en la ω. Este modo de proceder, al generalizarlo, conduce a decir que, si $f(\bar{x}, \bar{y})$ es bilineal simétrica, entonces $\omega(\bar{x}) = f(\bar{x}, \bar{x})$ es cuadrática.

DEFINICION

[116]

Sea V un espacio vectorial, sobre un cuerpo K. Se llaman *formas cuadráticas*, definidas en V, a las aplicaciones $\omega: V \to K$ que cumplen lo exigido en uno de los dos puntos siguientes (que son equivalentes entre sí):

I. Dada una forma bilineal simétrica $f: V \times V \to K$, se llama forma cuadrática asociada a f a la aplicación

$$\omega: V \to K, \quad \bar{x} \mapsto \omega(\bar{x}) = f(\bar{x}, \bar{x})$$

II. La aplicación $\omega: V \to K$ es una forma cuadrática si se verifica:

a) $\omega(\lambda\bar{x}) = \lambda^2\omega(\bar{x}), \quad \forall\, \bar{x} \in V \text{ y } \forall\, \lambda \in K$

b) La aplicación $f: V \times V \to K$ definida mediante la relación(*)

$$f(\bar{x}, \bar{y}) = \tfrac{1}{2}[\omega(\bar{x} + \bar{y}) - \omega(\bar{x}) - \omega(\bar{y})]$$

es bilineal (simétrica). A f se la llama *forma polar* de ω.

(*) Se supone que el cuerpo K tiene característica distinta de 2 (cosa que ocurre si $K = \mathbb{R}$ o si $K = \mathbb{C}$), es decir, si $1 + 1 \neq 0$, o sea $2 \neq 0$, lo que permite llegar a este relación tal como está, con el 1/2 que hay en su segundo miembro.

DEMOSTRACION

1. Empecemos comprobando que I $\Rightarrow$ II. De verificarse I, sería:

   a) $\omega(\lambda\bar{x}) = f(\lambda\bar{x}, \lambda\bar{x}) = \lambda f(\bar{x}, \lambda\bar{x}) = \lambda^2 f(\bar{x}, \bar{x}) = \lambda^2\omega(x)$

   b) $\omega(\bar{x}+\bar{y}) - \omega(x) - \omega(y) = f(\bar{x}+\bar{y}, \bar{x}+\bar{y}) - \omega(\bar{x}) - \omega(\bar{y}) =$
   $= [f(\bar{x}, \bar{x}) + f(\bar{x}, \bar{y}) + f(\bar{y}, \bar{x}) + f(\bar{y}, \bar{y})] - \omega(\bar{x}) - \omega(\bar{y}) =$
   $= \omega(x) + 2f(\bar{x}, \bar{y}) + \omega(\bar{y}) - \omega(\bar{x}) - \omega(\bar{y}) = 2f(\bar{x}, \bar{y})$

   Por tanto (como $2 \neq 0$) dividiendo por 2 se confirma lo dicho en el apartado *b)* de II.

2. Veamos ahora que, en el caso de cumplirse II, $\omega(\bar{x})$ es igual a $f(\bar{x}, \bar{x})$. De acuerdo con *b)* y tomando $\bar{x} = \bar{y}$, se obtiene que

$$f(\bar{x}, \bar{x}) = \tfrac{1}{2}\,[\omega(2\bar{x}) - 2\omega(\bar{x})]$$

   Ahora bien, de acuerdo con *a)*, es $\omega(2\bar{x}) = 4\omega(\bar{x})$ y llevando este resultado a la última igualdad, se obtiene que $f(\bar{x}, \bar{x}) = \omega(\bar{x})$, como había que comprobar.

EJEMPLOS

1. Si V es el espacio vectorial de las funciones continuas de [0, 1] en $\mathbb{R}$, la aplicación

$$f: V \times V \to \mathbb{R}, \quad (x(t), y(t)) \mapsto \int_0^1 x(t)y(t)\,dt$$

es una forma bilineal simétrica. La forma cuadrática asociada a f es la aplicación:

$$\omega: V \to \mathbb{R}, \quad x(t) \mapsto \int_0^1 x(t)^2\,dx$$

2. Sea $f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ la forma bilineal simétrica definida mediante la expresión:

$$f(\bar{x}, \bar{y}) = f((x_1, x_2), (y_1, y_2)) = 8x_1y_1 - 3x_1y_2 - 3x_2y_1 + 5x_2y_2 =$$
$$= [x_1\ x_2]\begin{bmatrix} 8 & -3 \\ -3 & 5 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = X^tAY$$

La forma bilineal asociada a f es la aplicación $\omega: \mathbb{R}^2 \to \mathbb{R}$ siguiente:

$$\omega(\bar{x}) = \omega(x_1, x_2) = f((x_1, x_2), (x_1, x_2)) = 8x_1^2 - 6x_1x_2 + 5x_2^2 =$$
$$= [x_1\ x_2]\begin{bmatrix} 8 & -3 \\ -3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = X^tAX$$

OBSERVACIONES

1. Si $\omega: V \to K$ es una forma cuadrática, entonces $\omega(\vec{o}) = 0$, lo que se prueba tomando $\lambda = 0$ en $\omega(\lambda\vec{x}) = \lambda\omega(\vec{x})$. Nótese que, en general, el recíproco puede no verificarse, o sea, $\omega(\vec{x}) = 0$ no implica $\vec{x} = \vec{o}$; así para $\omega(x, y) = x^2 - y^2$, es $\omega(1, 1) = 0$ aunque el vector $\vec{x} = (1, 1)$ no sea el vector nulo.
2. La expresión que proporciona la forma polar f en función de ω, esto es:

$$f(\vec{x}, \vec{y}) = \tfrac{1}{2}\,[\omega(\vec{x} + \vec{y}) - \omega(\vec{x}) - \omega(\vec{y})]$$

también se puede poner de la siguiente manera

$$f(\vec{x}, \vec{y}) = \tfrac{1}{4}\,[\omega(\vec{x} + \vec{y}) - \omega(\vec{x} - \vec{y})]$$

como se puede comprobar fácilmente.
3. Para definir una forma cuadrática ω puede recurrirse, también, a formas bilineales no simétricas. Si $g: V \times V \to K$ es una forma bilineal cualquiera, la aplicación $\omega: V \to K$ definida por $\omega(\vec{x}) = g(\vec{x}, \vec{x})$ es cuadrática; la forma polar de esta ω es la aplicación $f: V \times V \to K$ que viene dada por

$$\begin{aligned} f(\vec{x}, \vec{y}) &= \tfrac{1}{2}\,[\omega(\vec{x} + \vec{y}) - \omega(\vec{x}) - \omega(\vec{y})] = \\ &= \tfrac{1}{2}\,[g(\vec{x} + \vec{y}, \vec{x} + \vec{y}) - g(\vec{x}, \vec{x}) - g(\vec{y}, \vec{y})] = \\ &= \tfrac{1}{2}\,[g(\vec{x}, \vec{x}) + g(\vec{x}, \vec{y}) + g(\vec{y}, \vec{x}) + g(\vec{y}, \vec{y}) - g(\vec{x}, \vec{x}) - g(\vec{y}, \vec{y})] = \\ &= \tfrac{1}{2}\,[g(\vec{x}, \vec{y}) + g(\vec{y}, \vec{x})] \end{aligned}$$

4. Aplicación lineal asociada a una forma cuadrática. Dada una forma cuadrática $\omega: V \to K$ y si se llama f a su forma polar, para cada vector $\vec{u} \in V$ las aplicaciones parciales $f(_, \vec{u})$ y $f(\vec{u}, _)$ definidas por

$$f(_, \vec{u}): \vec{x} \mapsto f(\vec{x}, \vec{u}) \quad \text{y} \quad f(\vec{u}, _): \vec{x} \mapsto f(\vec{u}, \vec{x})$$

son aplicaciones lineales de V en K, según se dijo en [112], es decir, son formas lineales o elementos del espacio dual V^*. Como ahora f es simétrica, estas dos formas lineales $f(_, \vec{u})$ y $f(\vec{u}, _)$ son iguales. Pues bien, la aplicación ω_a (asociada a ω):

$$\omega_a: V \to V^*, \qquad \vec{u} \mapsto \omega_a(\vec{u}) = f(\vec{u}, _) = f(_, \vec{u})$$

es una aplicación lineal ya que, para cualesquiera $\vec{u}, \vec{v} \in V$ y $\lambda, \mu \in K$ es:

$$\omega_a(\lambda\vec{u} + \mu\vec{v}) = \lambda\omega_a(\vec{u}) + \mu\omega_a(\vec{v})$$

debido a que para cualquier vector $\bar{x} \in V$ es:

$$\omega_a(\lambda\bar{u} + \mu\bar{v})(\bar{x}) = f(\lambda\bar{u} + \mu\bar{v}, _)(\bar{x}) = f(\lambda\bar{u} + \mu\bar{v}, \bar{x}) =$$
$$= \lambda f(\bar{u}, \bar{x}) + \mu f(\bar{v}, \bar{x}) = \lambda f(\bar{u}, _)(\bar{x}) + \mu f(\bar{v}, _)(\bar{x}) =$$
$$= \lambda\omega_a(\bar{u})(\bar{x}) + \mu\omega_a(\bar{v})(\bar{x}) = (\lambda\omega_a(\bar{u}) + \mu\omega_a(\bar{v}))(\bar{x})$$

MATRIZ DE UNA FORMA CUADRATICA. CAMBIO DE BASE. RANGO

[117]

Sea V un espacio vectorial, sobre el cuerpo K, y sea $\omega: V \to K$ una forma cuadrática cuya forma polar es f. Si V tiene dimensión n y si $B = (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una base de V, entonces para cualquiera que sea $\bar{x} \in V$ es:

$$\omega(\bar{x}) = \sum_{i,j=1}^{n} a_{ij}x_ix_j = X^tAX = [x_1\ x_2\ \cdots\ x_n]\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix}$$

donde $A = [a_{ij}]$ es la matriz simétrica $n \times n$ en la que $a_{ij} = f(\bar{e}_i, \bar{e}_j)$, que se llama *matriz de la forma cuadrática* ω en la base B, y donde $X = [x_i]$ es la matriz columna de coordenadas de $\bar{x}$ en B.

Si A' es la matriz de ω respecto de otra base B', entonces

$$A' = P^tAP$$

donde P es la matriz (regular) del cambio de coordenadas(*) que se produce al cambiar la base B por la B'. Así, pues, A y A' son congruentes (véase [114]).

Todas las matrices asociadas a la forma cuadrática ω, en las distintas bases de V, tienen el mismo rango, al cual se le llama *rango de la forma cuadrática* ω:

$$\text{rang } \omega = \text{rang } A$$

(*) La ecuación del cambio de coordenadas es $X = PX'$, donde X y X' son las columnas de coordenadas, de un mismo vector, en las bases B y B', respectivamente.

DEMOSTRACION

Según sabemos, la expresión de $f(\bar{x}, \bar{y})$ es (véase [113]) $f(\bar{x}, \bar{y}) = X^tAY$; tomando aquí $\bar{x} = \bar{y}$, se obtiene que $\omega(\bar{x}) = X^tAY$.

Al cambiar de base, se verifica que (véase [114]) $f(\bar{x}, \bar{y}) = X''A'Y'$ donde $A' = P^tAP$; tomando $\bar{x} = \bar{y}$, se obtiene que $A' = P^tAP$ es la matriz de ω en la nueva base B'.

Como P es regular, A y A' son matrices equivalentes(*) y, por tanto, tienen el mismo rango. En consecuencia, la anterior definición de rango de ω es consistente.

OBSERVACION

Nótese que la matriz A es simétrica, por serlo la forma polar f, y que también lo son todas las matrices $A' = P^tAP$ congruentes con A.

Por ser A simétrica, esto es, como $a_{ij} = a_{ji}$ para todos los i y j, la expresión de $\omega(\bar{x})$ se puede expresar en la forma:

$$\omega(\bar{x}) = \sum_{i,j=1}^{n} a_{ij}x_ix_j = \sum_{h=1}^{n} a_{hh}x_h^2 + 2\sum_{i<j}^{n} a_{ij}x_ix_j$$

así, para $n = 3$, se tendría:

$$\omega(x) = (a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2) + 2(a_{12}x_1x_2 + a_{13}x_1x_3 + a_{33}x_2x_3)$$

EJERCICIO

De una forma cuadrática $\omega: \mathbb{R}^3 \to \mathbb{R}$ se sabe que

$$\omega(\bar{e}_1) = 0 \qquad \omega(\bar{e}_2) = -2 \qquad \omega(\bar{e}_3) = 3$$
$$\omega(\bar{e}_2 + \bar{e}_3) = 5 \quad \omega(\bar{e}_1 + \bar{e}_3) = 6 \qquad \omega(\bar{e}_1 + \bar{e}_2) = -4$$

(donde $\bar{e}_1$, $\bar{e}_2$ y $\bar{e}_3$ son los vectores de la base canónica).

Hallar la expresión analítica de ω.

RESOLUCION

Llamando f a la forma polar de ω, sabemos que

$$f(\bar{e}_1, \bar{e}_1) = \omega(\bar{e}_1) = 0$$
$$f(\bar{e}_2, \bar{e}_2) = \omega(\bar{e}_2) = -2$$
$$f(\bar{e}_3, \bar{e}_3) = \omega(\bar{e}_3) = 3$$
$$f(\bar{e}_1, \bar{e}_2) = \tfrac{1}{2}\,[\omega(\bar{e}_1 + \bar{e}_2) - \omega(\bar{e}_1) - \omega(\bar{e}_2)] = -1$$
$$f(\bar{e}_2, \bar{e}_3) = \tfrac{1}{2}\,[\omega(\bar{e}_2 + \bar{e}_3) - \omega(\bar{e}_2) - \omega(\bar{e}_3)] = 2$$
$$f(\bar{e}_1, \bar{e}_3) = \tfrac{1}{2}\,[\omega(\bar{e}_1 + \bar{e}_3) - \omega(\bar{e}_1) - \omega(\bar{e}_3)] = \tfrac{3}{2}$$

(*) Nótese que se puede poner $A' = Q^{-1}AP$, donde $Q^{-1} = P^t$.

Por tanto conocemos ya todos los elementos de la matriz simétrica de f, y de ω, en la base canónica, lo que permite poner:

$$\omega(x_1, x_2, x_3) = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 0 & -1 & 3/2 \\ -1 & -2 & 2 \\ 3/2 & 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$

$$= -2x_2^2 + 3x_3^2 - 2x_1x_2 + 3x_1x_3 + 4x_2x_3$$

ESPACIO VECTORIAL DE LAS FORMAS CUADRATICAS

[118]

Sea V un espacio vectorial, sobre el cuerpo K. El conjunto de las formas cuadráticas de V en K, al que denotaremos por $\mathcal{Q}(V)$, es un espacio vectorial, sobre K, respecto de las operaciones usuales (suma de funciones y producto de función por escalar).

La correspondencia que a cada forma cuadrática, de $\mathcal{Q}(V)$, le asigna su forma polar, forma bilineal simétrica de $\mathcal{B}_s(V)$, es un isomorfismo entre espacios vectoriales.

Si V tiene dimensión n y elegida una base en V, la correspondencia que a cada forma cuadrática, de $\mathcal{Q}(V)$, le asigna su matriz asociada, en esa base, es un isomorfismo del espacio vectorial $\mathcal{Q}(V)$ en el espacio vectorial de las matrices simétricas de tamaño $n \times n$.

DEMOSTRACION

1. Para comprobar que $\mathcal{Q}(V)$ es un espacio vectorial, bastará con comprobar que es un subespacio del espacio vectorial de las funciones de V en K, esto es, que $\mathcal{Q}(V)$ es no vacío (lo que es obvio, pues la aplicación nula es cuadrática) y que

$$[\omega, \omega' \in \mathcal{Q}(V) \text{ y } \lambda, \lambda' \in K] \Rightarrow \lambda\omega + \lambda'\omega' \in \mathcal{Q}(V)$$

Esto es así ya que, si f y f' son las formas polares de ω y ω', entonces $\lambda\omega + \lambda'\omega'$ es la forma cuadrática asociada a $\lambda f + \lambda' f'$ (que es bilineal y simétrica, ya que $\mathcal{B}_s(V)$ es espacio vectorial); esto es así ya que, para cualquier $\bar{x} \in V$ es:

$$(\lambda\omega + \lambda'\omega')(\bar{x}) = \lambda\omega(\bar{x}) + \lambda'\omega'(\bar{x}) = \lambda f(\bar{x}, \bar{x}) + \lambda' f'(\bar{x}, \bar{y}) =$$

$$= (\lambda f + \lambda' f')(\bar{x}, \bar{x})$$

2. La correspondencia que a cada forma bilineal simétrica $f \in \mathcal{B}_s(V)$ le asigna

su forma cuadrática asociada $\omega \in \mathfrak{Q}(V)$ es biyectiva (su recíproca asocia a cada forma cuadrática su forma polar); esta correspondencia, según ya sabemos (ver [116]), viene dada por:

$$\mathfrak{B}_s(V) \to \mathfrak{Q}(V), \quad f \mapsto \omega, \quad \text{siendo} \quad \omega(\bar{x}) = f(\bar{x}, \bar{x})$$

$$\mathfrak{Q}(V) \to \mathfrak{B}_s(V), \quad \omega \mapsto f, \quad \text{siendo} \quad f(\bar{x}, \bar{y}) = \tfrac{1}{2}\,[\omega(\bar{x}+\bar{y}) - \omega(\bar{x}) - \omega(\bar{y})]$$

Esta biyección es un isomorfismo entre espacios vectoriales ya que, según se acaba de comprobar en el apartado anterior, es lineal (para la suma y para el producto por escalar), es decir:

$$\begin{pmatrix} f = \text{forma polar de } \omega \\ f' = \text{forma polar de } \omega' \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda f + \lambda' f' \text{ es la forma} \\ \text{polar de } \lambda\omega + \lambda'\omega' \end{pmatrix}$$

3. La correspondencia que a cada forma cuadrática ω le asigna su matriz simétrica A, respecto de la base dada, es biyectiva de $\mathfrak{Q}(V)$ en el espacio de las matrices $n \times n$ simétricas. Nos estamos refiriendo a la biyección que ya conocemos, esto es a la dada (para cualquier $\bar{x} \in V$) por:

$$\omega(\bar{x}) = X'AX \quad (X = \text{columna de coordenadas de } \bar{x})$$

Esta biyección es un isomorfismo entre los referidos espacios, ya que si A, A' son las matrices de ω, $\omega' \in \mathfrak{Q}(V)$ y para cualesquiera λ, $\lambda' \in K$ es:

$$(\lambda\omega + \lambda'\omega')(x) = \lambda\omega(\bar{x}) + \lambda'\omega'(\bar{x}) = \lambda X'AX + \lambda' X'A'X = X'(\lambda A + \lambda' A')X$$

es decir, dicha biyección es lineal entre espacios vectoriales, o sea:

$$\begin{pmatrix} A = \text{matriz de } \omega \\ A' = \text{matriz de } \omega' \end{pmatrix} \Rightarrow \begin{pmatrix} \text{la matriz } \lambda A + \lambda' A' \text{ es} \\ \text{la asociada a } \lambda\omega + \lambda'\omega' \end{pmatrix}$$

7.3. CONJUGACION RESPECTO DE UNA FORMA CUADRATICA

Recuérdese que, cuando se estudiaba el producto escalar en $\mathbb{R}^3$, se dijo que dos vectores $\bar{x}$ e $\bar{y}$ eran ortogonales si su producto escalar era nulo. Pues bien, si se sustituye el producto escalar $(\bar{x}, \bar{y}) \mapsto \bar{x} \cdot \bar{y}$ por una forma bilineal cualquiera $(\bar{x}, \bar{y}) \mapsto f(\bar{x}, \bar{y})$, esto es, si en lugar de la relación $\bar{x} \cdot \bar{y} = 0$ se considera la $f(\bar{x}, \bar{y}) = 0$, se dice que $\bar{x}$ e $\bar{y}$ son conjugados, en vez de ortogonales. Esta generalización conserva muchas de las propiedades de la ortogonalidad de vectores, pero no todas. Así, por ejemplo, sabemos que un vector fijo $\bar{a} \neq \bar{o}$ no puede ser ortogonal a todos los vectores del espacio; sin embargo, en caso general, puede haber algún vector $\bar{a} \neq \bar{o}$ tal que $f(\bar{a}, \bar{x}) = 0$ para todo vector $\bar{x}$ del espacio.

VECTORES CONJUGADOS

[119]

Sea $\omega: V \to K$ una forma cuadrática, en el espacio vectorial V sobre el cuerpo K, y sea f la forma polar de ω. Se dice que $\bar{x}, \bar{y} \in V$ son *vectores conjugados* respecto de ω o respecto de f si se verifica que $f(\bar{x}, \bar{y}) = 0$.
Si V tiene dimensión finita y si A es la matriz simétrica de ω en una cierta base de V, entonces $\bar{x}, \bar{y} \in V$ son conjugados respecto de ω si y sólo si $X^t AY = O$ *(relación de conjugación)*, donde X e Y son las matrices columna de coordenadas de $\bar{x}$ e $\bar{y}$.

OBSERVACIONES

1. Como $f(\bar{x}, \bar{y}) = X^t AY$, la relación de conjugación se obtiene trivialmente anulando esta última expresión.
2. El vector nulo $\bar{o}$ es conjugado de todos los vectores del espacio, ya que $f(\bar{x}, \bar{o}) = 0$ para cualquiera que sea $\bar{x} \in V$.
3. La relación de conjugación respecto de ω es simétrica, esto es, si $\bar{x}$ es conjugado de $\bar{y}$ entonces $\bar{y}$ es conjugado de $\bar{x}$. Esto es evidente ya que $f(\bar{x}, \bar{y}) = f(\bar{y}, \bar{x})$, pues la forma polar f es simétrica.
4. Se dice que dos conjuntos de vectores (en particular, dos subespacios) son conjugados respecto de la forma cuadrática ω si cualquier vector de uno de ellos es conjugado de todos los vectores del otro.

EJEMPLO

Sea $\omega: \mathbb{R}^2 \to \mathbb{R}$ la forma cuadrática

$$\omega(\bar{x}) \equiv \omega(x_1, x_2) = x_1^2 + 3x_1x_2 + 2x_2^2$$

Los vectores $\bar{x}_0 = (2, -3)$ e $\bar{y}_0 = (6, -5)$ son conjugados respecto de ω ya que, como la forma polar de ω es:

$$f((x_1x_2), (y_1, y_2)) = x_1y_1 + \tfrac{3}{2}(x_1y_2 + x_2y_1) + 2x_2y_2$$

resulta que $f(\bar{x}_0, \bar{y}_0) = 0$.

Hallemos ahora todos los vectores $\bar{y} = (y_1, y_2)$ que son conjugados, respecto de ω, del vector $\bar{x}_0 = (2, -3)$. Para ello, poniendo $f(\bar{x}_0, \bar{y}) = 0$ se obtiene:

$$2y_1 + \tfrac{3}{2}(2y_2 - 3y_1) - 6y_2 = 0 \quad \text{o sea} \quad 5y_1 + 6y_2 = 0$$

Por tanto, todos los vectores conjugados de $\bar{x}_0$ forman el siguiente subespacio de $\mathbb{R}^2$:

$$\{(y_1, y_2) \in \mathbb{R}^2 / 5y_1 + 6y_2 = 0\}$$

A éste se le llama subespacio conjugado de $\bar{x}_0$; de estos subespacios vamos a hablar a continuación.

SUBESPACIO CONJUGADO DE UN VECTOR

[120]

Sea $\omega: V \to K$ una forma cuadrática, en el espacio vectorial V sobre el cuerpo K, y llamemos f a la forma polar de ω. Sea $\bar{x}_0 \in V$ un vector fijo. Se verifica que:

1. El conjunto que forman los vectores conjugados de $\bar{x}_0$, es decir,

$$\mathscr{C}(\bar{x}_0) = \{\bar{x} \in V / f(\bar{x}_0, \bar{x}) = 0\}$$

es un subespacio vectorial de V, se le llama *subespacio conjugado* del vector $\bar{x}_0$ respecto de ω.
2. Si $\omega(\bar{x}_0) \neq 0$, el subespacio $\mathscr{C}(\bar{x}_0)$ es suplementario del subespacio $\mathscr{V}(\bar{x}_0)$, que engendra $\bar{x}_0$.
3. Si $\omega(\bar{x}_0) \neq 0$ y si $(\bar{x}_0, \bar{x}_1, ..., \bar{x}_{n-1})$ es una base de V (se ha supuesto que dim $V = n$), entonces $\mathscr{C}(\bar{x}_0)$ admite por base a $(\bar{u}_1, ..., \bar{u}_{n-1})$ siendo

$$\bar{u}_i = \bar{x}_i - \lambda_i \bar{x}_0 \quad \text{con} \quad \lambda_i = \frac{f(\bar{x}_0, \bar{x}_i)}{\omega(\bar{x}_0)}$$

(para $i = 1, 2, ..., n-1$).

DEMOSTRACION

1. Hay que comprobar que, si $\bar{x}, \bar{x}' \in \mathscr{C}(\bar{x}_0)$ y $\lambda, \lambda' \in K$, entonces $\lambda\bar{x} + \lambda'\bar{x}'$ también pertenece a $\mathscr{C}(\bar{x}_0)$; así ocurre, ya que:

$$\bar{x}, \bar{x}' \in \mathscr{C}(\bar{x}_0) \Rightarrow f(\bar{x}_0, \bar{x}) = f(\bar{x}_0, \bar{x}') = 0 \Rightarrow f(x_0, \lambda\bar{x} + \lambda'\bar{x}') =$$
$$= \lambda f(\bar{x}_0, \bar{x}) + \lambda' f(\bar{x}_0, \bar{x}') = \lambda 0 + \lambda' 0 = 0 \Rightarrow \lambda\bar{x} + \lambda'\bar{x}' \in \mathscr{C}(\bar{x}_0)$$

2. Todo estriba en ver que $\mathscr{C}(\bar{x}_0) \oplus \mathscr{V}(\bar{x}_0) = V$, esto es, que:

- La suma de $\mathscr{C}(\bar{x}_0)$ y $\mathscr{V}(\bar{x}_0)$ es V. Esto es así debido a que todo $\bar{x} \in V$ se puede poner $\bar{x} = (\bar{x} - \lambda\bar{x}_0) + \lambda\bar{x}_0$, donde este último sumando pertenece a $\mathscr{V}(\bar{x}_0)$ para todo $\lambda \in K$ y el otro sumando, $\bar{x} - \lambda\bar{x}_0$, se puede conseguir que sea de $\mathscr{C}(\bar{x}_0)$ si se toma para λ el escalar que haga:

$$f(\bar{x}_0, \bar{x} - \lambda\bar{x}_0) = 0$$

o sea

$$f(\bar{x}_0, \bar{x}) - \lambda\omega(\bar{x}_0) = 0$$

es decir,

$$\lambda = \frac{f(\bar{x}_0, \bar{x})}{\omega(\bar{x}_0)}$$

- La intersección de $\mathscr{C}(\bar{x}_0)$ y $\mathscr{V}(\bar{x}_0)$ es nula ya que si $\bar{x}$ pertenece a $\mathscr{V}(\bar{x}_0)$ ha de ser $\bar{x} = \lambda\bar{x}_0$, para algún $\lambda \in K$, y para que $\bar{x}$ sea también de $\mathscr{C}(\bar{x}_0)$ ha de verificarse que $f(\bar{x}_0, \bar{x}) = 0$, luego tiene que ser:

$$f(\bar{x}_0, \lambda\bar{x}_0) = 0$$

esto es

$$\lambda\omega(\bar{x}_0) = 0$$

luego $\lambda = 0$ (pues $\omega(\bar{x}_0) \neq 0$). Por tanto $\bar{x} = 0\bar{x}_0 = \bar{o}$ es el único vector de la intersección.

3. Como $\mathscr{C}(\bar{x}_0) \oplus \mathscr{V}(\bar{x}_0) = V$ y las dimensiones de $\mathscr{V}(\bar{x}_0)$ y V son 1 y n, resulta que $\mathscr{C}(\bar{x}_0)$ tiene dimensión $n - 1$. Para comprobar que $(\bar{u}_1, ..., \bar{u}_{n-1})$ es una base de $\mathscr{C}(\bar{x}_0)$ debemos comprobar, pues, que los vectores $\bar{u}_i$ son todos de $\mathscr{C}(\bar{x}_0)$ y que forman un sistema independiente. Veamos que así ocurre:

- Los vectores $\bar{u}_i$ (para $i = 1, 2, ..., n - 1$) son de $\mathscr{C}(\bar{x}_0)$ ya que:

$$f(\bar{x}_0, \bar{u}_i) = f(\bar{x}_0, \bar{x}_i - \lambda_i\bar{x}_0) = f(\bar{x}_0, \bar{x}_i) - \lambda_i\omega(\bar{x}_0) =$$

$$= f(\bar{x}_0, \bar{x}_i) - \frac{f(\bar{x}_0, \bar{x}_i)}{\omega(\bar{x}_0)}\,\omega(\bar{x}_0) = 0$$

- El sistema $(\bar{u}_1, ..., \bar{u}_{n-1})$ es independiente ya que, si $h_1, ..., h_{n-1}$ son escalares, es:

$$h_1\bar{u}_1 + \cdots + h_{n-1}\bar{u}_{n-1} = \bar{o} \Rightarrow h_1(\bar{x}_1 - \lambda_1\bar{x}_0) + \cdots + h_{n-1}(\bar{x}_{n-1} - \lambda_{n-1}\bar{x}_0) = \bar{o} \Rightarrow$$

$$\Rightarrow h_1\bar{x}_1 + \cdots + h_{n-1}\bar{x}_{n-1} + (-h_1\lambda_1 - \cdots - h_{n-1}\lambda_{n-1})\bar{x}_0 = \bar{o} \overset{(*)}{\Rightarrow}$$

$$\Rightarrow [h_1 = 0, ..., h_{n-1} = 0, -h_1\lambda_1 - \cdots - h_{n-1}\lambda_{n-1} = 0] \Rightarrow h_1 = \cdots = h_{n-1} = 0$$

(la implicación con (*) es cierta debido a que $(\bar{x}_0, \bar{x}_1, ..., \bar{x}_{n-1})$ es base de V).

EJERCICIO

Sea $\omega: \mathbb{R}^3 \to \mathbb{R}$ la forma cuadrática definida, en la base canónica, mediante:

$$\omega(x, y, z) = x^2 - 3y^2 + z^2 + 4xy + 6xz = [x \quad y \quad z]\begin{bmatrix} 1 & 2 & 3 \\ 2 & -3 & 0 \\ 3 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = X^t A X$$

Hallar tres vectores $\bar{a}$, $\bar{b}$ y $\bar{c}$ conjugados dos a dos respecto de ω y tales que formen base de $\mathbb{R}^3$ (hay muchas soluciones de esta cuestión, se pide una de ellas). Hallar también la matriz de ω en la base $(\bar{a}, \bar{b}, \bar{c})$ que se obtenga.

RESOLUCION

Tomemos para $\bar{a}$ un vector cualquiera tal que $\omega(\bar{a}) \neq 0$; por ejemplo, $\bar{a} = (1, 0, 0)$. Los vectores $\bar{b}$ y $\bar{c}$ han de estar en el subespacio conjugado de $\bar{a}$, que es el:

$$[x \quad y \quad z]\, A \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = O \quad \text{o sea,} \quad x + 2y + 3z = 0 \qquad [1]$$

Tomemos para $\bar{b}$ un vector cualquiera de este subespacio; por ejemplo, $\bar{b} = (1, 1, -1)$. El vector $\bar{c}$, además de estar en el anterior subespacio, también ha de pertenecer al subespacio conjugado de $\bar{b}$, que es él:

$$[x \quad y \quad z] A \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = O \quad \text{o sea,} \quad -y + 2z = 0 \qquad [2]$$

Por tanto, el vector $\bar{c}$, que ha de verificar a [1] y a [2], puede ser el $\bar{c} = (-7, 2, 1)$.

En la base $(\bar{a}, \bar{b}, \bar{c})$, la matriz de ω será:

$$A' = \begin{bmatrix} \omega(\bar{a}) & f(\bar{a}, \bar{b}) & f(\bar{a}, \bar{c}) \\ f(\bar{b}, \bar{a}) & \omega(\bar{b}) & f(\bar{b}, \bar{c}) \\ f(\bar{c}, \bar{a}) & f(\bar{c}, \bar{b}) & \omega(\bar{c}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -60 \end{bmatrix}$$

FORMAS CUADRATICAS ORDINARIAS Y DEGENERADAS. NUCLEO

[121]

Sea $\omega: V \to K$ una forma cuadrática, en el espacio vectorial V sobre el cuerpo K. Se llama *núcleo de la forma cuadrática* ω al conjunto que forman los vectores de V que son conjugados, respecto de ω, de todos los vectores de V. Este conjunto es un subespacio de V y se le denota por:

$$\text{Nuc}\ (\omega) = \{\bar{x} \in V / f(\bar{x}, \bar{y}) = 0,\ \forall y \in V\}\ ^{(*)} \quad (f = \text{forma polar de } \omega)$$

Se dice que ω es una *forma cuadrática ordinaria* si Nuc $(\omega) = O$. Se dice que ω es una *forma cuadrática degenerada* si no es ordinaria, esto es, si existe algún vector $\bar{x}_0 \neq \bar{o}$ que es conjugado respecto de ω de todos los vectores de V.

Si dim $V = n$ y si es A la matriz de ω en una cierta base de V, entonces el núcleo de ω está formado por los $\bar{x} \in V$ tales que:

$$\text{Nuc}\ (\omega): AX = O \qquad (X = \text{columna de coordenadas de } \bar{x})$$

En este caso, de dimensión finita n, se verifica que:

$$\text{rang}\ \omega = n - \dim\ [\text{Nuc}\ (\omega)]$$
$$(\omega \text{ ordinaria}) \Leftrightarrow \text{rang}\ \omega = n$$
$$(\omega \text{ degenerada}) \Leftrightarrow \text{rang}\ \omega < n$$

(*) Para que un vector $\bar{x}$ sea del núcleo no basta con que sea $f(\bar{x}, \bar{x}) = 0$; cuando se verifica esta relación, es decir, si es $\omega(\bar{x}) = 0$, se dice que $\bar{x}$ es autoconjugado respecto de ω.

DEMOSTRACION

- El conjunto Nuc (ω) es, realmente, un subespacio de V ya que, para $\bar{x}_1, \bar{x}_2 \in V$ y $\lambda_1, \lambda_2 \in K$ es:

$$\bar{x}_1, \bar{x}_2 \in \text{Nuc}\ (\omega) \Rightarrow f(\bar{x}_1, \bar{y}) = f(\bar{x}_2, \bar{y}) = 0, \quad \forall \bar{y} \in V \Rightarrow$$
$$\Rightarrow f(\lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2, \bar{y}) = \lambda_1 f(\bar{x}_1, \bar{y}) + \lambda_2 f(\bar{x}_2, \bar{y}) = \lambda_1 0 + \lambda_2 0 = 0 \Rightarrow$$
$$\Rightarrow \lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2 \in \text{Nuc}\ (\omega)$$

- En el caso de dimensión finita es:

$$\bar{x} \in \text{Nuc}\ (\omega) \Leftrightarrow f(\bar{x}, \bar{y}) = 0,\ \forall\ \bar{y} \in V \Leftrightarrow$$
$$\Leftrightarrow Y^t AX = O,\ \forall Y \in \mathcal{M}_{n \times 1} \Leftrightarrow AX = O$$

lo que permite afirmar que $AX = O$ es la ecuación de Nuc (ω).

Como consecuencia de lo anterior; resulta que:

$$\dim [\text{Nuc } (\omega)] = \dim \{\text{soluciones de } AX = O\} =$$
$$= n - \text{rang } A = n - \text{rang } \omega$$
$$(\omega \text{ ordinaria}) \Leftrightarrow \text{Nuc } (\omega) = O \Leftrightarrow \dim [N(\omega)] = 0 \Leftrightarrow$$
$$\Leftrightarrow \text{rang } \omega = n$$
$$(\omega \text{ degenerada}) \Leftrightarrow (\omega \text{ no ordinaria}) \Leftrightarrow \text{rang } \omega \neq n \Leftrightarrow$$
$$\Leftrightarrow \text{rang } \omega < n$$

EJEMPLO

Sea $\omega: \mathbb{R}^3 \to \mathbb{R}$ la forma cuadrática que, respecto de la base canónica, tiene por expresión a:

$$\omega(x, y, z) = 3x^2 + y^2 + 4xy - 2xz - 2yz$$

El núcleo de ω lo forman los vectores (x, y, z) tales que:

$$\begin{bmatrix} 3 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

La matriz de este sistema (que es la matriz de ω en la base canónica) tiene rango 2, luego el núcleo de ω tiene dimensión $3 - 2 = 1$. Fácilmente se obtiene que la solución de este sistema es el conjunto de vectores $(x, y, z) = (\alpha, -\alpha, \alpha)$ para $\alpha \in \mathbb{R}$; por tanto:

$$\text{Nuc } (\omega) = \{(\alpha, -\alpha, \alpha) / \alpha \in \mathbb{R}\}$$
$$\text{rang } \omega = 2 \quad ; \quad \omega \text{ es degenerada}$$

OBSERVACION

Recuérdese que (ver la 4.ª observación que sigue a [116]), dada la forma cuadrática $\omega: V \to K$, a cuya forma polar llamaremos f, se llamaba aplicación asociada a ω a la

$$\omega_a: V \to V^* \quad , \quad \bar{u} \mapsto \omega_a(\bar{u}) = f(\bar{u}, _) = f(_, \bar{u})$$

que es lineal. El núcleo de ω_a está formado por aquellos vectores $\bar{u} \in V$ tales que $f(\bar{u}, _)$ es la aplicación nula, esto es, tales que $f(\bar{u}, \bar{y}) = 0$, para todo $\bar{y} \in V$, lo que equivale a decir que $\bar{u}$ pertenece al núcleo de ω. Así, pues, el núcleo de una forma cuadrática ω no es otra cosa que el núcleo de su aplicación lineal asociada ω_a.

β DIAGONALIZACION DE UNA FORMA CUADRATICA

Recuérdese que, en el caso de dimensión finita, la matriz simétrica A asociada a una forma cuadrática $\omega: V \to K$ varía al cambiar la base que se tome en V, pasando a ser la matriz $A' = P^t AP$ (donde P es la matriz regular de cambio de coordenadas), que es congruente con A. Las matrices congruentes con la matriz A son las matrices asociada a la forma cuadrática ω respecto de las distintas bases de V.

Vamos a ocuparnos aquí de elegir bases apropiadas en V de manera que, en ellas, las matrices asociadas a ω resulten especialmente sencillas; en concreto, se desea que dicha matriz sea diagonal. Veremos de aquí a poco que siempre hay alguna base en la que se consigue dicha diagonalización de ω. Dicho de otro modo, en la clase que forman todas las matrices congruentes con una matriz cuadrada, dada arbitrariamente, hay alguna de tales matrices que es diagonal.

7.4. FORMA DIAGONAL (CONGRUENCIA)

EXISTENCIA DE BASES DE VECTORES CONJUGADOS

[122] **Sea $\omega: V \to K$ una forma cuadrática, definida en un espacio vectorial V sobre un cuerpo K. Si V tiene dimensión finita, entonces existe alguna base de V formada por vectores conjugados, dos a dos respecto de ω.**

DEMOSTRACION

Llamemos $f: V \times V \to K$ la forma polar de ω.

En el caso particular $\omega = o$ (aplicación nula) la propiedad se verifica obviamente, pues dos vectores cualesquiera son conjugados respecto de $\omega = o$. Se supondrá pues que es $\omega \neq o$.

Demostremos la propiedad por inducción sobre $\dim V$. Si $\dim V = 1$, la propiedad es evidente. Hay, pues, que demostrar que si la propiedad es cierta para espacios de dimensión $n-1$, entonces también lo es si $\dim V = n$.

Sea $\bar{u}_1 \in V$ un vector tal que $\omega(\bar{u}_1) \neq 0$ (nótese que ha de ser $\bar{u}_1 \neq \bar{o}$), el cual existe pues se ha supuesto $\omega \neq o$, y recurramos al subespacio $V_1 = \mathscr{C}(\bar{u}_1)$ conjugado de $\bar{u}_1$ respecto de ω, que tiene dimensión $n-1$ y es suplementario del $\mathscr{V}(\bar{u}_1)$ (véase [120], 2). La aplicación $\omega: V_1 \to K$, restricción de ω a V_1, es una forma cuadrática cuya forma polar es $f: V_1 \times V_1 \to K$, restricción de f a $V_1 \times V_1$. Por la hipótesis de inducción, como $\dim V_1 = n-1$, en V_1 hay alguna

base $(\bar{u}_2, ..., \bar{u}_n)$ formada por vectores conjugados, dos a dos, respecto de ω. De todo ello resulta que para $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ se cumple lo que se exigía en el enunciado, pues:

- Es base de V, ya que $(\bar{u}_2, ..., \bar{u}_n)$ es base de V_1 y $V = V_1 \oplus \mathcal{V}(\bar{u}_1)$.
- Sus vectores son conjugados dos a dos, por serlo los $\bar{u}_2, ..., \bar{u}_n$ entre sí y, también, cada uno de ellos con $\bar{u}_1$, pues pertenecen al subespacio V_1 conjugado de $\bar{u}_1$.

DIAGONALIZACION DE UNA FORMA CUADRATICA

En el caso de dimensión finita, toda forma cuadrática ω es diagonalizable; esto se puede hacer de muchas formas, obteniéndose resultados distintos en unas y otras. Para ello se recurre a las bases que se acaban de estudiar; expresando ω en una tal base, su matriz asociada va a ser diagonal.

Al expresar este asunto en términos de matrices, se obtiene el siguiente resultado: de entre todas las matrices congruentes con una matriz simétrica dada, hay algunas que son diagonales. Veamos que las cosas son como acabamos de adelantar:

[123]

1. Si $\omega: V \to K$ es una forma cuadrática, definida en un espacio vectorial V de dimensión n, sobre el cuerpo K, entonces existe alguna base $B = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ de V en la que la matriz D asociada a ω es diagonal, o sea, tal que en dicha base es:

$$\omega(\bar{x}) = X^t D X = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2 \quad (*)$$

Diagonalizar una forma cuadrática ω es hallar una tal base B y la correspondiente matriz D.

2. El rango de ω es el número de elementos no nulos que hay en la diagonal de D. Suponiendo que los posibles d_i nulos ocupan los últimos lugares, es:

$$\omega(\bar{x}) = d_1 x_1^2 + \cdots + d_r x_r^2, \quad \text{donde} \quad \begin{cases} r \text{ es el rango de } \omega \\ d_i \neq 0 \quad \text{para} \quad i = 1, ..., r \end{cases}$$

3. Dada una matriz simétrica A cualquiera, de tamaño $n \times n$, existe alguna matriz regular P, de tamaño $n \times n$, tal que $D = P^t A P$ es una matriz diagonal. *Diagonalizar por congruencia* una matriz simétrica A es hallar las anteriores matrices D y P.

(*) X es la columna de coordenadas $x_1, x_2, ..., x_n$ de $\bar{x} \in V$ en la base B. Los elementos de la diagonal de D son $d_i = \omega(\bar{u}_i)$ para $i = 1, 2, ..., n$.

DEMOSTRACION

1. Según se acaba de comprobar (véase [121]) existe una base $B = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ de V formada por vectores conjugados, dos a dos respecto de ω, esto es, tal que $f(\bar{u}_i, \bar{u}_j) = 0$ para $i \neq j$ (con $i, j = 1, 2, ..., n$), donde f es la forma polar de ω. Como, por otra parte, la matriz de ω en la base B tiene por elemento de lugar ij al escalar $f(\bar{u}_i, \bar{u}_j)$, resulta que esta matriz tiene nulos todos sus elementos no situados en la diagonal principal ($i \neq j$), es decir, dicha matriz es diagonal.
2. El rango de ω es igual al rango de cualquiera de sus matrices asociadas y, en particular, rang ω = rang D. Como el rango de D es igual al número de elementos no nulos de su diagonal, la propiedad a demostrar es ya evidente.
3. Recurramos a la forma cuadrática $\omega: K^n \to K$ que tiene a A por matriz respecto de la base canónica. Según se acaba de decir en el punto anterior, hay alguna base B de K^n en la que la matriz de ω es diagonal; sea D esta matriz. Si es P la matriz del cambio de coordenadas que se produce al pasar de la base canónica a la base B, entonces (véase [117]) es $D = P^t AP$, como había que comprobar.

EJEMPLO

Considérese la forma cuadrática $\omega: \mathbb{R}^4 \to \mathbb{R}$ definida mediante

$$\omega(x_1, x_2, x_3, x_4) = [x_1 \quad x_2 \quad x_3 \quad x_4] \begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 3 & -2 & 1 \\ -1 & -2 & 9 & -5 \\ 1 & 1 & -5 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = X^t AX$$

Supongamos que se quiere diagonalizar ω de manera que la base en la que se obtenga la diagonalización sea una $\{\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4\}$ del tipo:

$$\bar{u}_1 = (x, 0, 0, 0) \quad \bar{u}_2 = (y, z, 0, 0) \quad \bar{u}_3 = (u, v, w, 0) \quad \bar{u}_4 = (\alpha, \beta, \gamma, \delta)$$

Tomemos $\bar{u}_1 = (1, 0, 0, 0)$. Como $\bar{u}_2$ ha de ser conjugado de $\bar{u}_1$, debe ser

$$[y \quad z \quad 0 \quad 0]A[1, 0, 0, 0]^t = O, \quad \text{o sea} \quad y + 2z = 0$$

Por tanto, tomamos $\bar{u}_2 = (2, -1, 0, 0)$. Como $\bar{u}_3$ ha de ser conjugado de $\bar{u}_1$ y de $\bar{u}_2$, ha de ser

$$\begin{matrix} [u & v & w & 0]A[1 & 0 & 0 & 0]^t = O \\ [u & v & w & 0]A[2 & -1 & 0 & 0]^t = O \end{matrix} \quad \text{o sea} \quad \begin{cases} u + 2v - w = 0 \\ v = 0 \end{cases}$$

Por tanto, tomamos $\bar{u}_3 = (1, 0, 1, 0)$. Como $\bar{u}_4$ ha de ser conjugado de $\bar{u}_1$, de $\bar{u}_2$ y de $\bar{u}_3$, ha de ser:

$$\left.\begin{array}{l}[\alpha\ \beta\ \gamma\ \delta]A[1\ 0\ 0\ 0]^t = O \\ [\alpha\ \beta\ \gamma\ \delta]A[2\ -1\ 0\ 0] = O \\ [\alpha\ \beta\ \gamma\ \delta]A[1\ 0\ 1\ 0] = O\end{array}\right\} \text{ o sea } \left\{\begin{array}{r} a + 2\beta - \gamma + \delta = 0 \\ \beta + \delta = 0 \\ 8\gamma - 4\delta = 0\end{array}\right.$$

Por tanto, tomamos $\bar{u}_4 = (3, -2, 1, 2)$.

Como $\omega(\bar{u}_1) = 1$, $\omega(\bar{u}_2) = -1$, $\omega(\bar{u}_3) = 8$ y $\omega(\bar{u}_4) = 6$, resulta que la matriz asociada a ω en la base $(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$ es

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

Nótese que esta diagonalización expresada matricialmente, $D = P^tAP$, sería:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 3 & -2 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 3 & -2 & 1 \\ -1 & -2 & 9 & -5 \\ 1 & 1 & -5 & 5 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & 0 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

Es conveniente observar aquí que los vectores $\bar{v}_i = a_i\bar{u}_i$, para cualesquiera $a_i \neq 0$ (para $i = 1, 2, 3, 4$), también son conjugados dos a dos y forman una base de $\mathbb{R}^4$ en la que la matriz de ω es diagonal y los elementos de su diagonal son $\omega(\bar{v}_i) = a_i^2\omega(\bar{u}_i)$. Tomando, entonces, $a_1 = 1$, $a_2 = 1$, $a_3 = \sqrt{8}$ y $a_4 = \sqrt{6}$, nos encontramos con que, en la base $(\bar{v}_i)$, la matriz de ω es diagonal y tiene su diagonal formada por los elementos 1, -1, 1 y 1.

7.5. DIAGONALIZACION EFECTIVA DE UNA FORMA CUADRATICA

Para diagonalizar por congruencia una matriz cuadrada A que sea simétrica, se puede recurrir a las operaciones elementales; cada operación elemental se realizará, primero, en las filas de A y, después, la misma operación, en las columnas de la matriz ya obtenida (se puede, también, actuar primero sobre las columnas y después sobre las filas; el resultado es el mismo). Procediendo de este modo se obtienen siempre matrices congruentes con A; entre ellas están, obviamente, las matrices diagonales que son congruentes con A. Más exactamente:

DIAGONALIZACION (POR CONGRUENCIA) USANDO OPERACIONES ELEMENTALES

[124]

1. Sea A una matriz cuadrada simétrica. Si E es una operación elemental, la matriz $E(A)$ que resulta de aplicar E a filas y a columnas en A (*) es congruente con A y tal que:

$$E(A) = FAC$$

donde,

$$\begin{cases} F = E^f(I) = \text{resultado de aplicar } E \text{ a las filas de } I \text{ (unidad)} \\ C = E^c(I) = \text{resultado de aplicar } E \text{ a las columnas de } I \\ (\text{Nótese que } C \text{ y } F \text{ son regulares y que } F = C^t) \end{cases}$$

2. Sea $\omega: V \to K$ una forma cuadrática, en un espacio vectorial V de dimensión finita sobre el cuerpo K, que tiene ecuación $\omega(\bar{x}) = X^tAX$, en una cierta base. Si E es una operación elemental para la que $D = E(A)$ es una matriz diagonal, entonces ω tiene expresión diagonal $\omega(\bar{x}) = X'^tDX'$ al realizar el cambio de coordenadas $X = CX'$.

(*) Primero se aplica E a las filas de A y, después, se aplica E a las columnas de la matriz ya obtenida. También se puede aplicar E primero a columnas y luego a filas.

COMPROBACION

1. Según ya sabemos (véase [026]), al aplicar a las filas de una matriz M una operación elemental E_1 y al aplicar a sus columnas una operación elemental E_2 se obtienen, respectivamente, las matrices

$$E_1^f(M) = FM \quad \text{con } F = E_1^f(I) \qquad \text{y} \qquad E_2^c(M) = MC \quad \text{con } C = E_2^c(I)$$

Por ello, la matriz $E(A)$ se puede poner en la forma:

$$E(A) = E^c(E^f(A)) = E^c(FA) = (FA)C = FAC$$
$$\text{con } F = E^f(I) \text{ y } C = E^c(I)$$

Es obvio y conocido que F y C son traspuesta la una de la otra. También sabemos (véase [028], 4) que F y C son matrices regulares. De esto resulta, pues, que

$$E(A) = C^tAC \quad \text{con } C \text{ regular,}$$

es decir, que $E(A)$ es congruente con A.

2. Si $D = E(A)$ es matriz diagonal, con la operación E se habrá diagonalizado por congruencia la matriz A, esta diagonalización será la dada por $D = C^tAC$. Por ello, con E se obtiene, como consecuencia, la diagonalización de la forma cuadrática ω; esta diagonalización se consigue, pues, al cambiar de base en V de manera que sea C la matriz de cambio de coordenadas.

DIAGONALIZACION EFECTIVA DE UNA FORMA CUADRATICA

Realizando adecuadas operaciones elementales, sobre filas y sobre columnas, siempre es posible diagonalizar por congruencia una matriz A simétrica dada. Esto es, aplicando unas ciertas operaciones elementales, en las filas y en las columnas de su matriz simétrica, toda forma cuadrática (en el caso de dimensión finita) es diagonalizable:

[125]

Para cualquiera que sea la matriz simétrica A, existe alguna operación elemental E tal que la matriz $D = E(A)$ (*) es diagonal. Como $E(A)$ es congruente con A, con una tal operación elemental se diagonaliza por congruencia a la matriz A; dicha diagonalización es:

$$D = P^tAP$$

donde

$$D = E(A) \quad ; \quad P = E^c(I) \quad \text{y} \quad P^t = E^f(I) \quad (*)$$

Utilizando como pivotes a los elementos 11, 22, ..., sucesivamente, y realizando la misma operación en filas y en columnas a A (matriz simétrica), se van obteniendo sucesivamente matrices en cuyas 1.ª, 2.ª, ..., fila y columna son nulos todos los elementos no ubicados en la diagonal. Si alguno de dichos elementos ii fuese nulo, entonces previamente, a las fila y columna i-ésima o se las suma otras posteriores, o se las permuta con ellas, de modo que, con esto, resulta un elemento 11, 22, ..., no nulo. Procediendo de este modo, a la postre se acaba obteniendo una matriz diagonal $D = E(A)$.

Si, al tiempo que se hace lo dicho con A, se aplican las mismas operaciones a las filas (sólo a las filas; no a las filas y columnas) de la matriz unidad I, se obtiene simultáneamente $P^t = E^f(I)$.

(*) $E(A)$ es la matriz si se obtiene aplicando E primero a las filas de A y, después, a las columnas de la matriz que se haya obtenido. Las matrices $E^f(I)$ y $E^c(I)$ se obtienen de aplicar E o las filas y columnas, respectivamente, de la matriz unidad I.

COMPROBACION

Veamos que el método descrito conduce, realmente, a la diagonalización por congruencia de A (simétrica) y veámoslo por inducción sobre el tamaño n de A. Para $n = 1$ la matriz A es diagonal, por lo que, sin necesidad de hacer nada, el método funciona. Suponiendo que el método funciona para matrices simétricas de tamaño $(n-1) \times (n-1)$, hay que comprobar que también funciona para las de tamaño $n \times n$. Sea, pues, A una matriz simétrica $n \times n$.

- Supongamos primero que el elemento de lugar 11 de A es no nulo. Utilizando a este elemento como pivote, es obvio que, con las oportunas operaciones elementales sobre filas, se transforman en ceros todos los elementos situados, en la primera columna, por debajo del 11. Como A es simétrica, si ahora se realizan las mismas operaciones elementales sobre las columnas de la matriz obtenida, se llega a una matriz simétrica A' que tiene nulos, excepto el elemento 11, todos los demás de sus primera fila y primera columna. La matriz A_1 que resulta de prescindir, en A', de estas primera fila y primera columna es simétrica y de tamaño $(n-1) \times (n-1)$. De acuerdo con nuestra hipótesis de inducción, esta matriz A_1 se puede diagonalizar realizando ciertas operaciones elementales en filas y, las mismas, en columnas. Haciendo estas manipulaciones en la matriz completa A', es evidente que se obtiene la diagonalización por congruencia de la matriz simétrica dada A.
- Supongamos ahora que el elemento de lugar 11 es nulo. Si no fuese nulo alguno de los elementos ii, podemos empezar con el intercambio $1.^a \rightleftarrows i^a$ (en filas y columnas), con lo que se obtiene una matriz simétrica cuyo elemento de lugar 11 es no nulo (pues es el anterior elemento ii); procediendo ahora con esta matriz del modo indicado en el punto anterior, se obtiene la diagonalización buscada. Si todos los elementos ii son nulos y no lo es alguno de los $i1$, podemos empezar con la operación $1.^a \rightarrow 1.^a + i^a$ (en filas y columnas), con lo que se obtiene una matriz simétrica cuyo elemento de lugar 11 no es nulo (ya que es igual al doble del anterior $1i$); procediendo ahora con esta matriz del modo indicado en el punto anterior, se obtiene la diagonalización buscada.

  Finalmente, si todos los elementos $1i$ son nulos, la diagonalización por congruencia de A se reduce a la de la matriz simétrica de tamaño $(n-1) \times (n-1)$ que resulta de prescindir de sus primera fila y primera columna.

EJEMPLO

Sea $\omega: \mathbb{R}^4 \rightarrow \mathbb{R}$ la forma cuadrática que respecto de la base canónica $(\bar{e}_1, \bar{e}_2, \bar{e}_3, \bar{e}_4)$ tiene asociada la siguiente matriz simétrica A:

$$A = \begin{bmatrix} 0 & 2 & 6 & 0 \\ 2 & 1 & -1 & 2 \\ 6 & -1 & 3 & -3 \\ 0 & 2 & -3 & 5 \end{bmatrix}$$

Para diagonalizar ω, diagonalicemos A por congruencia realizando en sus filas y columnas las siguientes operaciones elementales, que también realizaremos en las filas de la matriz unidad:

Primera $(1.^a) \rightleftarrows (2.^a)$

Segunda $(2.^a) \to (2.^a) - 2(1.^a)$, $(3.^a) \to (3.^a) + (1.^a)$ y $(4.^a) \to (4.^a) - 2(1.^a)$

Tercera $(3.^a) \to (3.^a) + 2(2.^a)$ y $(4.^a) \to (4.^a) - (2.^a)$

Cuarta $(4.^a) \to 2(4.^a) + (3.^a)$

Procediendo de este modo se obtiene sucesivamente:

$$\left[\begin{array}{cccc} 1 & 2 & -1 & 2 \\ 2 & 0 & 6 & 0 \\ -1 & 6 & 3 & -3 \\ 2 & 0 & -3 & 5 \end{array}\right]\left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & -4 & 8 & -4 \\ 0 & 8 & 2 & -1 \\ 0 & -4 & -1 & 1 \end{array}\right]\left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 18 & -9 \\ 0 & 0 & -9 & 5 \end{array}\right]\left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 2 & -3 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right]\left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 2 & -3 & 0 & 1 \\ 0 & -3 & 1 & 2 \end{array}\right]$$

Por tanto, hemos obtenido la siguiente expresión diagonal de ω:

$$\omega(\bar{x}) = x_1^2 - 4x_2^2 + 18x_3^2 + 2x_4^2$$

donde x_1, x_2, x_3, x_4 son las coordenadas del vector $\bar{x}$ en la siguiente nueva base $(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$:

$$\begin{aligned} \bar{u}_1 &= \bar{e}_2 \\ \bar{u}_2 &= \bar{e}_1 - 2\bar{e}_2 \\ \bar{u}_3 &= 2\bar{e}_1 - 3\bar{e}_2 + \bar{e}_4 \\ \bar{u}_4 &= -3\bar{e}_4 + \bar{e}_3 + 2\bar{e}_4 \end{aligned}$$

OBSERVACION

Sea $\omega: V \to K$ una forma cuadrática, donde V es un espacio vectorial de dimensión n sobre el cuerpo K. Supongamos que en cierta base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V la expresión de $\omega(\bar{x})$ para $\bar{x} \in V$ es:

$$\omega(\bar{x}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = X^t A X$$

Sea $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ una base en la que ω tiene expresión diagonal, esto es, tal que

$$\omega(\bar{x}) = d_1 x_1'^2 + d_2 x_2'^2 + \cdots + d_n x_n'^2 = X'^t D X'$$

Sea P la matriz del correspondiente cambio de coordenadas, de manera que $X = PX'$. Denotando a la matriz P^{-1} en la forma que a continuación se indica, el cambio de coordenadas, $X' = P^{-1}X$, se puede poner:

$$P^{-1} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ l_1 & l_2 & \cdots & l_n \end{bmatrix} \qquad \begin{aligned} x_1' &= a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \\ x_2' &= b_1 x_1 + b_2 x_2 + \cdots + b_n x_n \\ &\cdots\cdots\cdots \\ &\cdots\cdots\cdots \\ x_n' &= l_1 x_1 + l_2 x_2 + \cdots + l_n x_n \end{aligned}$$

Por tanto, la expresión diagonal de $\omega(\bar{x})$ nos permite poner:

$$\begin{aligned} \omega(\bar{x}) = {} & d_1(a_1 x_1 + a_2 x_2 + \cdots + a_n x_n)^2 + \\ & + d_2(b_1 x_1 + b_2 x_2 + \cdots + b_n x_n)^2 + \\ & \cdots\cdots\cdots \\ & \cdots\cdots\cdots \\ & + d_n(l_1 x_1 + l_2 x_2 + \cdots + l_n x_n)^2 \end{aligned} \qquad \text{(I)}$$

Obsérvese que el resultado de todo esto se puede resumir diciendo que la expresión $\omega(\bar{x}) = \sum a_{ij} x_i x_j$ se ha puesto como suma de constantes por cuadrados de expresiones lineales de las variables; obsérvese que estas expresiones, o formas, lineales son independientes entre sí por ser P regular. Este modo (I) de escribir $\omega(\bar{x})$ recibe el nombre de «descomposición en cuadrados» de la forma cuadrática.

γ FORMAS CUADRATICAS REALES

Se llaman formas cuadráticas reales a aquéllas que están definidas en un espacio vectorial real; así pues, aquí los escalares van a ser los números reales. Por ello, vamos a poder hablar ahora de «escalares positivos» y de «escalares

negativos», cosa que no se puede hacer, en general, con un cuerpo de escalares cualquiera.

En este apartado nos ocuparemos de aquellas cuestiones, relativas a una forma cuadrática real $\omega: V \to \mathbb{R}$, en las que interviene la relación de orden $\leqslant$ entre números reales y, en particular, con el signo que toma la imagen $\omega(\bar{x})$, de un vector $\bar{x} \in V$.

7.6. FORMAS DEFINIDAS Y LEY DE INERCIA

FORMAS CUADRATICAS DEFINIDAS Y SEMIDEFINIDAS

[126]

Sea V un espacio vectorial real y considérese la forma cuadrática real $\omega: V \to \mathbb{R}$; sea f la forma polar de ω. Se dice que ω o que f es:

$$\textit{definida}^{(*)} \begin{cases} \text{positiva si } [\omega(\bar{x}) > 0 \text{ para todo } \bar{x} \neq \bar{o} \text{ de } V] \\ \text{negativa si } [\omega(\bar{x}) < 0 \text{ para todo } \bar{x} \neq \bar{o} \text{ de } V] \end{cases}$$

$$\textit{semidefinida} \begin{cases} \text{positiva si } [\omega(\bar{x}) \geqslant 0,\ \forall \bar{x} \in V, \text{ y } \omega \text{ no es definida}] \\ \text{negativa si } [\omega(\bar{x}) \leqslant 0,\ \forall \bar{x} \in V, \text{ y } \omega \text{ no es definida}] \end{cases}$$

Si dim $V = n$, si D es una matriz diagonal asociada a ω en una cierta base (que sabemos que existe) y llamando $d_1, d_2, \ldots, d_n$ a los elementos de la diagonal de D, se verifica que:

1. ω es definida $\begin{cases} \text{positiva} \Leftrightarrow [d_i > 0,\ i = 1, 2, \ldots, n] \\ \text{negativa} \Leftrightarrow [d_i < 0,\ i = 1, 2, \ldots, n] \end{cases}$

2. ω semidefinida $\begin{cases} \text{positiva} \Leftrightarrow [d_i \geqslant 0,\ \forall i \text{ y } d_{i_0} = 0 \text{ para algún } i_0] \\ \text{negativa} \Leftrightarrow [d_i \leqslant 0,\ \forall i \text{ y } d_{i_0} = 0 \text{ para algún } i_0] \end{cases}$

(*) Aun cuando los términos «definida positiva» y «definida negativa» son los utilizados, quizá fuese preferible sustituirlos por «positivamente definida» y «negativamente definida» o, también, por «estrictamente positiva» y «estrictamnte negativa».

DEMOSTRACION

Sean $x_1, x_2, \ldots, x_n$ las coordenadas de un vector $\bar{x}$ en la base del enunciado; a ésta la llamaremos $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n)$. Con dicha notación, es:

$$\omega(\bar{x}) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2$$

1. Consideremos el caso de ω definida positiva (para ω definida negativa se procede de modo análogo).

   - Si ω es definida positiva, entonces ha de ser $d_i > 0$ para $i = 1, 2, ..., n$, ya que si para cierto i fuese $d_i \leqslant 0$, tomando $\bar{x} = \bar{u}_i$ sería $\omega(\bar{x}) = d_i \leqslant 0$, que no es posible ya que es $\bar{x} \neq \bar{o}$.
   - Si $d_i > 0$ para $i = 1, 2, ..., n$, entonces ω es definida positiva. En efecto, dado $\bar{x} \neq \bar{o}$, como alguna de sus coordenadas no es nula, por ejemplo, $x_i \neq 0$, resulta que:

$$\omega(\bar{x}) = d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2 \geqslant d_ix_i^2 > 0$$

2. Consideramos el caso de ω semidefinida positiva (para ω semidefinida negativa se razona análogamente).

   - Si ω es semidefinida positiva, entonces ha de ser $d_i \geqslant 0$ para todo i, pues si fuese $d_i < 0$ para algún i, tomando $\bar{x} = \bar{u}_i$ sería $\omega(\bar{x}) = d_i < 0$, que no es posible. Además, tiene que ser $d_{i_0} = 0$ para algún i_0 ya que de no ser así ω sería definida.
   - Si $d_i \geqslant 0$ para todo i y, además, $d_{i_0} = 0$. entonces ω ha de ser semidefinida positiva. En efecto, para todo $\bar{x} \in V$, los n sumandos de la expresión de $\omega(\bar{x})$ son no negativos, luego $\omega(\bar{x}) \geqslant 0$; por otra parte para $\bar{x} = \bar{u}_{i_0}$ será $\omega(\bar{x}) = d_{i_0}x_{i_0}^2 = 0$.

EJEMPLOS

1. La forma cuadrática $\omega: \mathbb{R}^3 \to \mathbb{R}$ dada por

$$\omega(x, y, z) = x^2 + (x - y)^2 + (x + y - z)^2 = 3x^2 + 2y^2 + z^2 - 2xz - 2yz$$

es definida positiva ya que:

$$\omega(x, y, z) = 0 \Leftrightarrow [x = 0, x - y = 0, x + y - z = 0] \Leftrightarrow x = y = z = 0$$

2. Si V es el espacio vectorial de las funciones integrables de $[0, 1]$ en $\mathbb{R}$, la forma cuadrática

$$\omega: V \to \mathbb{R}, \quad f \mapsto \omega(f) = \int_0^1 f(x)^2\,dx$$

es semidefinida positiva ya que, como $f(x)^2 \geqslant 0$ en $[0, 1]$, es:

$$\int_0^1 f(x)\,dx \geqslant 0$$

y, además, $\omega(f) = 0$ para alguna función f no nula, como la siguiente:

$$f(x) = 0 \quad \text{para} \quad x \neq \tfrac{1}{2} \qquad \text{y} \qquad f(\tfrac{1}{2}) = 1$$

EJERCICIO

Hallar los valores de $\alpha \in \mathbb{R}$ para los que la siguiente forma cuadrática $\omega: \mathbb{R}^3 \to \mathbb{R}$ es definida positiva o semidefinida positiva:

$$\omega(x, y, z) = x^2 + 2y^2 + \alpha z^2 + 2xy + 2\alpha yz$$

RESOLUCION

Diagonalicemos ω por congruencia (sólo nos interesa una matriz diagonal asociada a ω; no tiene aquí utilidad hallar la correspondiente base de $\mathbb{R}^3$); realizando operaciones elementales en filas y en columnas de la matriz de ω, se obtiene sucesivamente:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & \alpha \\ 0 & \alpha & \alpha \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & \alpha & \alpha \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha - \alpha^2 \end{bmatrix}$$

Por tanto, en una cierta base, en las que las coordenadas de un vector $\bar{x} \in \mathbb{R}$ se denotan por x_1, x_2, x_3, es:

$$\omega(\bar{x}) = x_1^2 + x_2^2 + \alpha(\alpha - 1)x_3^2$$

Resulta entonces que

$$\begin{aligned} \alpha < 0 \text{ o } \alpha > 1 &\Rightarrow \alpha - \alpha^2 < 0 \Rightarrow \omega \text{ no es definida ni semidefinida} \\ \alpha = 0 \text{ o } \alpha = 1 &\Rightarrow \alpha - \alpha^2 = 0 \Rightarrow \omega \text{ es semidefinida positiva} \\ 0 < \alpha < 1 &\Rightarrow \alpha - \alpha^2 > 0 \Rightarrow \omega \text{ es definida positiva} \end{aligned}$$

OBSERVACIONES

- Hay muchas formas cuadráticas que no son ni definidas ni semidefinidas, como ha quedado de manifiesto en el ejercicio anterior. No estamos, pues, clasificando las formas cuadráticas; solamente hemos destacado algunas de ellas (las definidas y semidefinidas), que tienen especial relevancia.
- Hay quienes llaman formas cuadráticas positivas a las que son definidas positivas o semidefinidas positivas y, también, formas cuadráticas negativas a las que son definidas negativas o semidefinidas negativas. Esto es, dada una forma cuadrática real $\omega: V \to \mathbb{R}$, se suele definir:

$$\begin{aligned} [\omega \text{ positiva}] &\Leftrightarrow [\omega(\bar{x}) \geqslant 0, \ \forall \bar{x} \in V] \\ [\omega \text{ negativa}] &\Leftrightarrow [\omega(\bar{x}) \leqslant 0, \ \forall \bar{x} \in V] \end{aligned}$$

- Si $\omega: V \to \mathbb{R}$ es una forma cuadrática definida positiva, entonces $-\omega$ es una forma cuadrática definida negativa y recíprocamente.

- Sea $\omega: V \to \mathbb{R}$ una forma cuadrática que, en cierta base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V, tiene asociada la matriz $A = [a_{ij}]$. Si ω es definida positiva, entonces los elementos $a_{11}, a_{22}, ..., a_{nn}$ son todos positivos, ya que $a_{ii} = \omega(\bar{e}_i) > 0$ para $i = 1, 2, ..., n$. Sin embargo, no es suficiente con que todos los a_{ii} sean positivos para poder asegurar que ω es definida positiva. Así, por ejemplo, la forma cuadrática $\omega: \mathbb{R}^2 \to \mathbb{R}$ dada por

$$\omega(x, y) = x^2 + y^2 + 3xy$$

tiene positivos los coeficientes de $x^2 + y^2$ pero no es definida positiva pues, por ejemplo, $\omega(1, -2) = -1 < 0$.

MATRICES SIMETRICAS DEFINIDAS

- Se dice que una matriz real simétrica A de tamaño $n \times n$ es definida positiva si $X^tAX > 0$ para toda matriz columna $X \neq O$, de tamaño $n \times 1$. Es decir, A es definida positiva si es matriz de una forma cuadrática definida positiva.

De análoga manera se definen las matrices simétricas definidas negativas y semidefinidas positivas y negativas. Esto es:

[127]

Matrices definidas y semidefinidas. Se dice que una matriz real y simétrica A, de tamaño $n \times n$, es definida positiva, definida negativa, semidefinida positiva o semidefinida negativa si lo es, respectivamente, la forma cuadrática $\omega: \mathbb{R}^n \to \mathbb{R}$ asociada a la matriz A en la base canónica.

Si D es una matriz diagonal congruente con A (que sabemos que existe) y llamando $d_1, d_2, ..., d_n$ a los elementos de la diagonal de D, se verifica que:

1. A es definida $\begin{cases} \text{positiva} & \Leftrightarrow [d_i > 0,\ i = 1, 2, ..., n] \\ \text{negativa} & \Leftrightarrow [d_i < 0,\ i = 1, 2, ..., n] \end{cases}$

2. A semidefinida $\begin{cases} \text{positiva} & \Leftrightarrow [d_i \geqslant 0,\ \forall i \text{ y } d_{i_0} = 0 \text{ para algún } i_0] \\ \text{negativa} & \Leftrightarrow [d_i \leqslant 0,\ \forall i \text{ y } d_{i_0} = 0 \text{ para algún } i_0] \end{cases}$

DEMOSTRACION

La matriz D está asociada a la forma cuadrática ω en una determinada base de $\mathbb{R}^n$; por tanto, según lo anteriormente demostrado, los apartados 1) y 2) se verifican si se sustituye A por ω. Como en esto de ser definida o semidefinida, lo que le ocurra a ω le ocurre también a A (por definición), resulta que los apartados 1) y 2) son ciertos (tal y como están enunciados, esto es, para la matriz A).

LEY DE INERCIA DE SYLVESTER. SIGNATURA

Dada una forma cuadrática ω, definida en un espacio vectorial V de dimensión n, hay bases de V en las que (véase [123]) la expresión de $\omega(\bar{x})$, para $\bar{x} \in V$, es:

$$\omega(\bar{x}) = d_1x_1^2 + d_2x_2^2 + \cdots + d_rx_r^2$$

$$\text{donde} \quad \begin{cases} d_1 \neq 0,\ d_2 \neq 0,\ \ldots,\ d_r \neq 0 \\ r \text{ es el rango de } \omega \\ x_i = \text{coordenada } i\text{-ésima de } \bar{x} \end{cases}$$

En nuestro caso, cuando ω es una forma cuadrática real, estos d_i (no nulos) serán positivos o negativos y, cambiando si es necesario el orden de los vectores de la base, se puede suponer que los positivos ocupan los primeros lugares y los negatios los últimos. Llamemos p el número de los d_i que son positivos, que los podremos denotar poniendo $d_i = a_i^2$; los restantes d_i serán negativos y los podremos denotar poniendo $d_i = -a_i^2$. Con ello, resulta que $\omega(\bar{x})$ se puede poner:

$$\omega(\bar{x}) = [(a_1x_1)^2 + \cdots + (a_px_p)^2] - [(a_{p+1}x_{p+1})^2 + \ldots + (a_rx_r)^2]$$

Pues bien, se verifica que el número p de sumandos positivos que hay en la anterior expresión es el mismo para todas las bases en las que $\omega(\bar{x})$ adopta forma diagonal. Veamos esto con detalle:

[128]

Sea V un espacio vectorial real de dimensión finita y $\omega: V \to \mathbb{R}$ una forma cuadrática (real). Los números p y q de elementos positivos y negativos de cualquiera de las matrices diagonales asociadas a ω, en las correspondientes bases de V, son invariantes, los mismos para todas las diagonalizaciones; esto es, los números p y q sólo dependen de ω. Se llama *signatura* de ω al par:

$$\text{sig } (\omega) = (p, q) \qquad \text{(nótese que } p + r = \text{rang } \omega)$$

DEMOSTRACION

Sean $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n)$ y $(\bar{u}'_1, \bar{u}'_2, \ldots, \bar{u}'_n)$ dos bases de V en las que ω tiene asociadas matrices diagonales; sean p y p' el número de elementos positivos de una y otra matriz. Hemos de probar que $p = p'$.

Para ello recurramos a los siguientes subespacios U y W de V:

$$U = \mathcal{V}(\bar{u}_1, \ldots, \bar{u}_p) \quad \text{y} \quad W = \mathcal{V}(\bar{u}'_{p'+1}, \ldots, \bar{u}'_n)$$

Nótese que la restricción de ω a U es definida positiva pues tiene asociada, en la base $(\bar{u}_1, \ldots, \bar{u}_p)$, una matriz diagonal con los elementos de su diagonal

todos positivos. También hay que notar que la restricción de ω a W es semidefinida negativa pues tiene asociada, en la base $(\vec{u}'_{p'+1}, ..., \vec{u}'_n)$, una matriz diagonal sin elementos positivos.

Resulta entonces que $U \cap W = O$, ya que si $\vec{x} \neq \vec{o}$ fuese de $U \cap W$, por ser de U se verificaría que $\omega(\vec{x}) > 0$ y por ser de W se verificaría que $\omega(\vec{x}) \leq 0$ y estas dos desigualdades son incompatibles. Por tanto, como $U + W \subset V$, será

$$n = \dim V \geq \dim (U + W) = \dim U + \dim W - \dim U \cap W = p + (n - p') + 0$$

es decir $n \geq n + (p - p')$, lo que equivale a que sea $p \leq p'$. Intercambiando los papeles de p y p' y repitiendo así lo hasta aquí dicho, se obtiene que también ha de ser $p' \leq p$. Por consiguiente, se verifica que $p = p'$, como había que comprobar.

SIGNATURA DE UNA MATRIZ SIMETRICA

[129]

Se llama *signatura* de una matriz real y simétrica A de tamaño $n \times n$ a la signatura de la forma cuadrática $\omega: \mathbb{R}^n \to \mathbb{R}$ asociada a la matriz A en la base canónica. Se verifica que:

1. La signatura de una matriz real y simétrica A es sig $A = (p, q)$, donde p y q son los números de elementos positivos y negativos de cualquier matriz diagonal congruente con A.
2. La signatura de una matriz real y simétrica A es invariante por congruencia, es decir, es igual al de cualquier matriz $A' = P^t AP$ (donde P es regular), congruente con A.

DEMOSTRACION

1. Las matrices diagonales congruentes con A son matrices diagonales asociadas, todas ellas, a la forma cuadrática ω, en las correspondientes bases. Por tanto, según se acaba de probar en [128], todas estas matrices tienen los mismos números p y q de elementos positivos y negativos y el par (p, q) es la signatura de ω, luego dicho par es la signatura de A.
2. Como las matrices diagonales que son congruentes con A' son las mismas que son congruentes con A, y de acuerdo con la propiedad precedente, resulta que A' y A tienen el mismo rango.

EJEMPLO

Hallemos la signatura de la siguiente matriz simétrica A:

$$A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & 3 & 1 & -2 \\ 0 & 1 & -5 & -2 \\ -1 & -2 & -2 & 0 \end{bmatrix}$$

Para ello, diagonalicemos A por congruencia; esto lo haremos realizando las siguientes operaciones elementales (en filas y en columnas):

Primera (2.ª) → (2.ª) − 2(1.ª) y (3.ª) → (3.ª) + (1.ª)
Segunda (3.ª) → (3.ª) + (2.ª)
Tercera (4.ª) → (4.ª) − (1/2)(3.ª)

Procediendo de este modo se obtiene sucesivamente:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -5 & -2 \\ 0 & 0 & -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -4 & -2 \\ 0 & 0 & -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = D$$

Por tanto, resulta que:

$$\left.\begin{array}{l} [\text{número de elementos positivos de } D] = 1 \\ [\text{número de elementos negativos de } D] = 2 \end{array}\right\} \Rightarrow \operatorname{sig} A = (1, 2)$$

7.7. EXPRESION CANONICA DE UNA FORMA CUADRATICA

PROPOSICION

[130]

Sea $\omega: V \to \mathbb{R}$ una forma cuadrática, en el espacio vectorial real V de dimensión n. Si $\operatorname{sig} \omega = (p, q)$, con lo que $r = p + q$ es el rango de ω, entonces existe alguna base de V en la que la matriz de ω es la siguiente matriz diagonal C, que llamaremos matriz canónica de ω:

$$C = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & -1 & \\ & & & & & & 0 \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{bmatrix} \begin{matrix} \\ p \\ \\ \\ q \\ \\ \\ n-r \\ \end{matrix}$$

$\operatorname{sig} \omega = (p, q)$
$\operatorname{rang} \omega = p + q = r$
$\dim V = n$

Si $(x_1, x_2, \ldots, x_n)$ son las coordenadas de un vector $\bar{x} \in V$ en la citada base, entonces la expresión de $\omega(\bar{x})$ es:

$$\omega(\bar{x}) = [x_1^2 + \cdots + x_p^2] - [x_{p+1}^2 + \cdots + x_r^2] \quad (\textit{expresión canónica})$$

DEMOSTRACION

Como la signatura de ω es (p, q), resulta que cualquiera de las matrices diagonales asociadas a ω tiene p elementos positivos y q elementos negativos. Sea $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ una de las bases en las que $\omega(\bar{x})$ adopta expresión diagonal; entonces, para ciertos escalares $d_1, ..., d_r$ positivos, es:

$$\omega(\bar{x}) = [d_1 x_1'^2 + \cdots + d_p x_p'^2] - [d_{p+1} x_{p+1}'^2 + \cdots + d_r x_r'^2]$$

donde $(x_1', x_2', ..., x_n')$ son las coordenadas de $\bar{x}$ en $(\bar{u}_2, \bar{u}_2, ..., \bar{u}_n)$. Es evidente que, como $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ es una base de V, también forman base los vectores $\bar{e}_1, \bar{e}_2, ..., \bar{e}_n$ siguientes:

$$\bar{e}_i = \begin{cases} \dfrac{1}{\sqrt{d_i}}\,\bar{u}_i & \text{para } i = 1, 2, ..., r \quad \text{(recuérdese que } d_i > 0\text{)} \\ \bar{u}_i & \text{para } i = r+1, ..., n \end{cases}$$

Con este cambio de base, las nuevas coordenadas $(x_1, x_2, ..., x_n)$ y las antiguas están relacionadas mediante:

$$x_i' = \frac{1}{\sqrt{d_i}}\, x_i \quad \text{para } i = 1, 2, ..., r$$

$$x_i' = x_i \quad \text{para } i = r+1, ..., n$$

lo que llevado a la anterior expresión de $\omega(\bar{x})$, conduce a

$$\omega(\bar{x}) = [x_1^2 + \cdots + x_p^2] - [x_{p+1}^2 + \cdots + x_r^2]$$

Así pues, en la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ la matriz de ω es la canónica.

PROPOSICION

[131]

Dos matrices reales y simétricas A y A', del mismo tamaño, son congruentes si, y sólo si, sig A = sig A'. Si la signatura de A vale sig $A = (p, q)$, entonces A es congruente con la matriz diagonal C cuya diagonal es $\overset{p}{1}, ..., 1, \overset{q}{-1}, ..., -1, 0, ..., 0$ (p «unos», q «menos unos» y el resto «ceros»); a C se la llama *matriz canónica de congruencia* de A.

DEMOSTRACION

- Si A y A' son congruentes entonces están asociadas a una misma forma cuadrática ω, en unas ciertas bases. Por tanto, sig A = sig ω y sig A' = sig ω, de donde resulta que sig A = sig A'.

- Si A y A' tienen la misma signatura, que llamaremos (p, q), entonces ambas son congruentes con la anterior matriz canónica C y, por ello, son congruentes entre sí.

OBSERVACION

Recuérdese que la congruencia de matrices es una relación de equivalencia, para el conjunto $\mathcal{S}_n$ de las matrices cuadradas simétricas de tamaño $n \times n$. Todas las matrices de $\mathcal{S}_n$ que son congruentes entre sí forman una clase de equivalencia; cada clase está formada por todas las matrices que tienen una misma signatura. Como representante de la clase de matrices que tienen signatura (p, q) se tomará la matriz canónica C (diagonal) cuya diagonal es

$$\overbrace{1, \ldots, 1}^{p}, \overbrace{-1, \ldots, -1}^{q}, 0, \ldots, 0$$

El número de clases de la relación de congruencia, entre matrices simétricas de tamaño $n \times n$, es entonces $(n+1)(n+2):2$.

CARACTERIZACION DE LAS FORMAS CUADRATICAS DEFINIDAS Y SEMIDEFINIDAS

[132]

Sea $\omega: V \to \mathbb{R}$ una forma cuadrática, en un espacio vectorial real V de dimensión n. Se verifica que:

1. ω es definida positiva $\Leftrightarrow$ sig $\omega = (n, 0)$.
2. ω es definida negativa $\Leftrightarrow$ sig $\omega = (0, n)$.
3. ω semidefinida positiva $\Leftrightarrow$ sig $\omega = (r, 0)$, con $r < n$.
4. ω semidefinida negativa $\Leftrightarrow$ sig $\omega = (0, r)$, con $r < n$.

DEMOSTRACION

Sea D una matriz diagonal asociada a ω, en una cierta base, y llamemos d_1, d_2, ..., d_n a los elementos de la diagonal de D. De acuerdo con lo ya dicho (véase [126]) sobre las formas cuadráticas definidas y semidefinidas, se tiene (sólo se considera el caso «positivo»; para el «negativo» se razona de igual modo):

- La forma ω es definida positiva si, y sólo si, todos los d_i son positivos, lo que equivale a que sea sig $\omega = (n, 0)$.
- La forma ω es semidefinida positiva si, y sólo si, ningún d_i es negativo y alguno de ellos es nulo, lo que equivale a que sea sig $\omega = (r, 0)$, donde $r < n$.

CARACTERIZACION DE LAS MATRICES SIMETRICAS DEFINIDAS

[133]

Sea A una matriz real y simétrica, de tamaño $n \times n$. Llamemos $I_{p,q}$ a la matriz diagonal de tamaño $n \times n$ cuya diagonal principal es $\overset{p}{1}, \ldots, 1, \overset{p}{-1}, \ldots, -1, 0, \ldots, 0$ (p «unos», q «menos unos» y el resto «ceros»). Se verifica que:

1. A es definida positiva $\Leftrightarrow$ A es congruente con $I_{n,0} = I$ (unidad).
2. A es definida negativa $\Leftrightarrow$ A es congruente con $I_{0,n} = -I$.
3. A semidefinida positiva $\Leftrightarrow$ A es congruente con $I_{r,0}$, siendo $r < n$.
4. A semidefinida negativa $\Leftrightarrow$ A es congruente con $I_{0,r}$, siendo $r < n$.

DEMOSTRACION

Sea $\omega : \mathbb{R}^n \to \mathbb{R}$ la forma cuadrática asociada a la matriz A en la base canónica. De acuerdo con la definición de matrices simétricas definidas y semidefinidas (véase [127]), se tiene (sólo se considera el caso «positivo»; para el caso «negativo» se razona de igual modo):

- A es definida positiva si, y sólo si, lo es ω; esto equivale a que la signatura de ω sea $(n, 0)$, lo que significa que $I_{n,0}$ es la matriz canónica de ω, es decir, que A es congruente con $I_{n,0}$.
- A es semidefinida positiva si, y sólo si, lo es ω; esto equivale a que sea sig $\omega = (r, 0)$ para cierto $r < n$, lo que significa que $I_{r,0}$ es la matriz canónica de ω, es decir, que A es congruente con $I_{r,0}$ (donde $r < n$).

OBSERVACION

Según la anterior caracterización de las formas cuadráticas reales $\omega : V \to \mathbb{R}$ (donde dim $V = n$) definidas y semidefinidas (véase [132]), las expresiones canónicas de estas formas son (aquí $x_1, x_2, \ldots, x_n$ son las coordenadas de un vector $\bar{x} \in V$, en la base en la que se obtiene la expresión canónica):

$$\omega \text{ es definida positiva} \to \omega(\bar{x}) = x_1^2 + x_2^2 + \cdots + x_n^2$$

$$\omega \text{ es definida negativa} \to \omega(\bar{x}) = -x_1^2 - x_2^2 - \cdots - x_n^2$$

$$\omega \text{ semidefinida positiva} \to \omega(\bar{x}) = x_1^2 + x_2^2 + \cdots + x_r^2 \text{ (con } r < n)$$

$$\omega \text{ semidefinida negativa} \to \omega(\bar{x}) = -x_1^2 - x_2^2 - \cdots - x_r^2 \text{ (con } r < n)$$

EJEMPLO

Obtengamos la expresión canónica, hallando también la correspondiente base $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$, de la forma cuadrática $\omega: \mathbb{R}^3 \to \mathbb{R}$ que en la base canónica tiene por expresión a

$$\omega(x, y, z) = x^2 + y^2 + z^2 + 4xy - 4xz + 4yz$$

Para ello, diagonalicemos por congruencia la matriz de ω, que llamaremos A, hasta llegar a la matriz canónica de congruencia. Esto se consigue con las siguientes operaciones elementales (en filas y en columnas):

primera $(2.^a) \to (2.^a) - 2(1.^a)$ y $(3.^a) \to (3.^a) + 2(1.^a)$

segunda $(3.^a) \to (3.^a) + 2(2.^a)$

tercera $(2.^a) \to (2.^a) : \sqrt{3}$ y $(3.^a) \to (3.^a) : 3$

cuarta $(2.^a) \rightleftarrows (3.^a)$

Con ello, se obtiene:

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 2 & -2 & 1 & 0 & 0 \\ 2 & 1 & 2 & 0 & 1 & 0 \\ -2 & 2 & 1 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -3 & 6 & -2 & 1 & 0 \\ 0 & 6 & -3 & 2 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & -2 & 1 & 0 \\ 0 & 0 & 9 & -2 & 2 & 1 \end{array}\right] \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2/\sqrt{3} & 1/\sqrt{3} & 0 \\ 0 & 0 & 1 & -2/3 & 2/3 & 1/3 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2/3 & 2/3 & 1/3 \\ 0 & 0 & -1 & -2/\sqrt{3} & 1/\sqrt{3} & 0 \end{array}\right]$$

$$\omega(\bar{x}) = x_1^2 + x_2^2 - x_3^2 \qquad \begin{array}{l} \bar{u}_1 = (1, 0, 0) \\ \bar{u}_2 = (-2/3, 2/3, 1/3) \\ \bar{u}_3 = (-2/\sqrt{3}, 1/\sqrt{3}, 0) \end{array}$$

CRITERIO DE SYLVESTER

Para averiguar si una forma cuadrática es definida positiva o definida negativa, se dispone del siguiente criterio de carácter práctico:

[134]

Sea $\omega: V \to \mathbb{R}$ una forma cuadrática, en el espacio vectorial real V de dimensión n. Si $A = [a_{ij}]$ es la matriz de ω (en una base cualquiera de V), sean $\Delta_1, \Delta_2, ..., \Delta_n$ los determinantes $\Delta_i = \det A_i$, donde

$$A_i = \begin{bmatrix} a_{11} & \cdots & a_{1i} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{ii} \end{bmatrix}$$

se verifica que:

1. [ω definida positiva] $\Leftrightarrow$ [$\Delta_i > 0$ para $i = 1, 2, ..., n$].
2. [ω definida negativa] $\Leftrightarrow$ [$(-1)^i \Delta_i > 0$ para $i = 1, 2, ..., n$].

DEMOSTRACION

Sea $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ la base en la que A es matriz de ω; sean $(x_1, x_2, ..., x_n)$ las coordenadas de un $\bar{x} \in V$ respecto de dicha base.

a) [Implicación $\Rightarrow$ del Apartado 1] Si ω es definida positiva, entonces también lo es su restricción a cualquier subespacio de V y, en particular, lo es su restricción $\omega_i = \omega|_{V_i}$ al subespacio $V_i = \mathcal{V}(\bar{u}_1, ..., \bar{u}_i)$ para $i = 1, 2, ..., n$. La matriz de ω_i en la base $(\bar{u}_1, ..., \bar{u}_i)$ de V_i es la matriz A_i. Como ω_i es definida positiva, su matriz es congruente con la matriz unidad I (de tamaño $i \times i$), es decir, hay una matriz regular P (de tamaño $i \times i$) tal que $A_i = P^t I P = P^t P$. Tomando determinantes en esta igualdad, se obtiene que $\Delta_i = (\det P)^2 > 0$ (para $i = 1, 2, ..., n$), como había que comprobar.

b) [Implicación $\Leftarrow$ del Apartado 1] Demostremos esta implicación por inducción sobre n. La implicación es evidente para $n = 1$. Suponiendo, pues, que la implicación es cierta para las formas cuadráticas en espacios de dimensión $n - 1$, hemos de probar que también lo es en los de dimensión n. Consideremos la restricción $\omega' = \omega|_U$, de ω al subespacio $U = \mathcal{V}(\bar{u}_1, ..., \bar{u}_{n-1})$. Nótese que ω' es una forma cuadrática en un espacio de dimensión $n - 1$ cuya matriz asociada en la base $(\bar{u}_1, ..., \bar{u}_{n-1})$ es la matriz A_{n-1} y que, por ello, los determinantes de la hipótesis son, para ω', los $\Delta_1 > 0$, $\Delta_2 > 0$, ..., $\Delta_{n-1} > 0$. Así pues, según la hipótesis de inducción, ω' es definida positiva y, por ello, existe una base $(\bar{e}_1, ..., \bar{e}_{n-1})$ de U en la que la matriz de ω' es la unidad, de tamaño $(n - 1) \times (n - 1)$. Por tanto, la matriz de ω en la base $(\bar{e}_1, ..., \bar{e}_{n-1}, \bar{u}_n)$ de V es del tipo

$$A' = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & c_{n-1} \\ \hline c_1 & c_2 & \cdots & c_{n-1} & c_n \end{array}\right] \quad \begin{pmatrix} \text{para ciertos escalares} \\ c_1, c_2, \ldots, c_{n-1}, c_n \in \mathbb{R} \end{pmatrix}$$

También será matriz de ω, en la base que corresponda, aquella que se obtiene de realizar en A' la siguiente operación en filas y en columnas:

$$(n\text{-ésima}) \rightarrow (n\text{-ésima}) - [c_1(1.^a) + c_2(2.^a) + \cdots + c_{n-1}(n-1)^a]$$

Haciendo esto se obtiene la siguiente matriz A'' (que también está asociada a ω por ser congruente con A'):

$$A'' = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \hline 0 & 0 & \cdots & 0 & d \end{array}\right] \quad \begin{pmatrix} \text{donde } d \text{ es el escalar} \\ d = c_n - (c_1^2 + c_2^2 + \cdots + c_{n-1}^2) \end{pmatrix}$$

Como A y A'' están ambas asociadas a ω, son congruentes y, por ello, para cierta matriz regular P es $A'' = P^t A P$: tomando determinantes en esta igualdad, se obtiene:

$$d = \det A'' = (\det P)^2 (\det A) = (\det P)^2 \Delta_n > 0$$

Como ha resultado ser $d > 0$, nos encontramos con que sig $A'' = (n, 0)$, luego sig $\omega = (n, 0)$, es decir, ω es definida positiva, como había que comprobar.

c) [Apartado 2] La forma ω es definida negativa si, y sólo si, $-\omega$ es definida positiva. Como la matriz de $-\omega$ es $-A = [-a_{ij}]$, de acuerdo con el apartado 1, resulta que $-\omega$ es definida positiva si, y sólo si, son positivos los determinantes de las matrices $-A_i$ (para $i = 1, 2, \ldots, n$). Dado que

$$\det(-A_i) = (-1)^i \det A_i = (-1)^i \Delta_i$$

de todo lo dicho se concluye que ω es definida negativa si, y sólo si, $(-1)^i \Delta_i > 0$ para $i = 1, 2, \ldots, n$, como había que comprobar.

CAPITULO

8

Espacios vectoriales euclídeos

Los «vectores geométricos» que ya conocemos de antiguo, aquellos que llamábamos «segmentos orientados», tenían bastantes más atributos que los que tienen los que venimos manejando últimamente. Bien es verdad que las propiedades de éstos son reflejo de las de aquéllos, pero también es cierto que las nociones de longitud (distancia) y de ángulo entre vectores se han quedado en el camino que, al generalizar, nos llevó desde los vectores libres del espacio ordinario a los espacios vectoriales abstractos.

Ya es momento de ocuparnos de estas nociones olvidadas y ello lo haremos incorporando, a nuestro espacio vectorial general, una nueva urdimbre algebraica en la que sustentar los referidos conceptos de distancia y ángulo. Nos estamos refiriendo al «producto escalar». Si dotamos, al espacio vectorial, de un producto escalar, vamos a poder trasladar, al terreno de lo abstracto, las nociones de longitud y ángulo; esto se hace de modo que se conserven las propiedades que a ellas les son propias.

¿Cómo introducir el concepto de producto escalar de manera que las cosas sigan funcionando, en lo esencial, como en el caso de los segmentos orientados? Aquí hay que actuar, como en todo proceso de abstracción, mirando con atención al producto escalar concreto, entre vectores geométricos, el cual nos es (o nos debiera ser) muy familiar, y extraer, con lucidez, sus propiedades esenciales, aquellas que lo caracterizan, de las que se deducen los demás. Estas propiedades, enunciadas para el caso general, serán nuestro punto de partida, esto es, serán los axiomas del producto escalar.

Y ya, para acabar esta introducción, hagamos una observación de interés: los escalares no pueden ser aquí elementos de un cuerpo cualquiera, han de ser números reales o complejos. Para nosotros el caso de interés es el caso real; no nos ocuparemos del producto escalar complejo. En resumen, en este capítulo el cuerpo de escalares será $K = \mathbb{R}$.

8

α PRODUCTO ESCALAR

8.1. PRODUCTO ESCALAR DE VECTORES

DEFINICION

[135]

Sea V un espacio vectorial real y considérese una aplicación que a cada par de vectores $(\bar{x}, \bar{y})$ de V le asigna un número real, que se denotará poniendo $\bar{x} \cdot \bar{y}$. Se dice que una tal aplicación

$$V \times V \to \mathbb{R} \quad , \quad (x, y) \mapsto \bar{x} \cdot \bar{y}$$

es un *producto escalar*, o un producto interno, si para cualesquiera que sean $\bar{x}, \bar{x}', \bar{y} \in V$ y $\lambda, \lambda' \in \mathbb{R}$ se verifica que:

1. $\bar{x} \cdot \bar{y} = \bar{y} \cdot \bar{x}$.
2. $(\lambda \bar{x} + \lambda' \bar{x}') \cdot \bar{y} = \lambda \bar{x} \cdot \bar{y} + \lambda' \bar{x}' \cdot \bar{y}$.
3. $\bar{x} \cdot \bar{x} > 0$ para todo $\bar{x} \neq \bar{o}$.

Dicho de otro modo, un producto escalar en V es cualquier aplicación, de $V \times V$ en $\mathbb{R}$, que sea bilineal, simétrica y definida positiva.
Se llama *espacio vectorial euclídeo* a todo espacio vectorial real dotado de un producto escalar.

EJEMPLOS

1. Citemos, en primer lugar, el ejemplo que nos ha servido para introducir este tema. El producto escalar (usual) entre vectores libres del espacio verifica a la anterior axiomática. Recuérdese que el producto escalar $\bar{u} \cdot \bar{v}$, de dos vectores libres $\bar{u}$ y $\bar{v}$, se definía como el producto de sus módulos por el coseno del ángulo que forman, es decir:

$$\bar{u} \cdot \bar{v} = |\bar{u}|\,|\bar{v}| \cos(\bar{u}, \bar{v})$$

2. Entre los vectores de $\mathbb{R}^n$, según se comprueba con facilidad, la relación

$$(x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

define un producto escalar, que llamaremos *producto escalar canónico*; con él, $\mathbb{R}^n$ es pues un espacio vectorial euclídeo, que se llama *espacio vectorial euclídeo usual* $\mathbb{R}^n$.

OBSERVACIONES

1. De los axiomas se deduce que, en un espacio vectorial euclídeo V, para $\bar{x}, \bar{y}, \bar{y}' \in V$ y $\lambda, \lambda' \in \mathbb{R}$, se verifica que:
   - $\bar{x} \cdot (\lambda\bar{y} + \lambda'\bar{y}') = \lambda\bar{x} \cdot \bar{y} + \lambda'\bar{x} \cdot \bar{y}'$.
   - $\bar{x} \cdot \bar{o} = 0$.

   Estas propiedades se demuestran fácilmente; quedan, pues, como ejercicio.
2. El tercero de los axiomas del producto escalar se puede enunciar también diciendo que:

$$\bar{x} \cdot \bar{x} \geqslant 0 \text{ para todo } \bar{x} \in V \quad \text{y} \quad \bar{x} \cdot \bar{x} = 0 \Leftrightarrow \bar{x} = \bar{o}$$

   La equivalencia de ésta y la anterior forma de enunciar este axioma es clara; téngase en cuenta a este respecto que $\bar{o} \cdot \bar{o} = 0$.
3. Para denotar el producto escalar de dos vectores $\bar{x}$ e $\bar{y}$, además de $\bar{x} \cdot \bar{y}$ se suele poner también $\langle \bar{x}, \bar{y} \rangle$ ó $(\bar{x} \mid \bar{y})$. Cuando se precise indicar cuál es el producto escalar que se está manejando en V, al correspondiente espacio euclídeo se le representa poniendo $(V, \cdot)$, $(V, \langle , \rangle)$ o $(V, (\mid))$, respectivamente.
4. Aún cuando aquí no precisemos de ello, bueno es conocer que en el caso complejo el primer axioma del producto escalar se sustituye por $\bar{y} \cdot \bar{x} = \overline{\bar{x} \cdot \bar{y}}$, donde $\overline{\bar{x} \cdot \bar{y}}$ es el complejo conjugado de $\bar{x} \cdot \bar{y}$. Sin esta precaución, las axiomas serían incompatibles, pues nos encontraríamos con que, siendo $\bar{x} \cdot \bar{x} > 0$ para todo $\bar{x} \neq \bar{o}$, si tomamos $\bar{y} = i\bar{x}$, sería $\bar{y} \cdot \bar{y} = (i\bar{x}) \cdot (i\bar{x}) = (i^2)\bar{x} \cdot \bar{x} = -\bar{x} \cdot \bar{x} < 0$ para un cierto $\bar{y} \neq \bar{o}$, lo que va contra el axioma 3.

EJERCICIOS

1. En el espacio vectorial $V = \mathscr{C}(I, \mathbb{R})$, de las funciones reales que son continuas en un intervalo compacto $I = [a, b]$, se puede definir el producto escalar de dos funciones $f, g \in V$ mediante

$$(f \mid g) = \int_a^b f(x)g(x)\,dx$$

Veamos que para esta definición se cumple la axiomática del producto escalar:

- $(f \mid g) = \displaystyle\int_a^b f(x)g(x)\,dx = \int_a^b g(x)f(x)\,dx = (g \mid f)$

- $(\lambda f + \lambda' f' \mid g) = \displaystyle\int_a^b (\lambda f(x) + \lambda' f'(x))g(x)\,dx = \lambda \int_a^b f(x)g(x)\,dx +$

$$+ \lambda' \int_a^b f'(x)g'(x)\,dx = \lambda(f \mid g) + \lambda'(f' \mid g)$$

- $(f \mid f) = \int_a^b f(x)^2\,dx > 0$ si $f \neq o$ (ya que la integral de una función continua no negativa, como $f(x)^2$, que no es la función nula, es estrictamente positiva).

2. En el espacio vectorial $V = \mathcal{M}_n$ de las matrices reales de tamaño $n \times n$, se puede definir un producto escalar de dos matrices $A = [a_{ij}]$ y $B = [b_{ij}]$ mediante

$$\langle A, B\rangle = \text{traza}\,(AB^t) = \sum_{i,j=1}^{n} a_{ij}b_{ij}$$

veamos que para esta forma de definir $\langle A, B\rangle$ se verifican los tres axiomas del producto escalar:

- $\langle A, B\rangle = \sum_{i,j} a_{ij}b_{ij} = \sum_{i,j} b_{ij}a_{ij} = \langle B, A\rangle$.
- $\langle \lambda A + \lambda' A', B\rangle = \sum_{i,j} (\lambda a_{ij} + \lambda' a'_{ij})b_{ij} = \lambda_i \sum_{i,j} a_{ij}b_{ij} + \lambda' \Sigma\, a'_{ij}b_{ij} =$
$$= \lambda\langle A, B\rangle + \lambda'\langle A', B\rangle.$$
- $\langle A, A\rangle = \sum_{i,j} a_{ij}a_{ij} = \sum_{i,j} a_{ij}^2 > 0$ salvo si $A = [a_{ij}] = O$.

EXPRESION DEL PRODUCTO ESCALAR

[136]

Sea V un espacio vectorial euclídeo(*). Si $\dim V = n$, si $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una base de V y si se conocen los n^2 productos escalares $\bar{e}_i \cdot \bar{e}_j$ (para $i, j = 1, 2, ..., n$), entonces el producto escalar $\bar{x} \cdot \bar{y}$, de dos vectores cualesquiera $\bar{x}, \bar{y} \in V$, es igual a:

$$\bar{x} \cdot \bar{y} = \sum_{i,j=1}^{n} g_{ij}x_i y_j = X^t G Y \quad \text{siendo} \quad \begin{cases} g_{ij} = \bar{e}_i \cdot \bar{e}_j \ (i, j = 1, 2, ..., n) \\ G = [g_{ij}] \text{ matriz } n \times n \end{cases}$$

donde $(x_1, ..., x_n)$ e $(y_1, ..., y_n)$ son las coordenadas de $\bar{x}$ e $\bar{y}$ y X e Y son sus columnas de coordenadas. La matriz $G = [g_{ij}]$, que se llama *matriz métrica* (o de Gram) del producto escalar en la base dada, es una matriz simétrica definida positiva (véase [127] y [134]).

Al realizar un cambio de base, llamando P a la matriz del correspondiente cambio de coordenadas ($X = PX'$), la matriz métrica en la nueva base es la

$$G' = P^t G P \quad (G \text{ y } G' \text{ congruentes})$$

(*) Salvo indicación en contra, el producto escalar de dos vectores $\bar{x}$ e $\bar{y}$ de V se denotará poniendo $\bar{x} \cdot \bar{y}$. No obstante, en lugar de llamar V al espacio vectorial euclídeo, formalmente debiéramos designarle escribiendo $(V, \cdot)$.

DEMOSTRACION

Si bien es cierto que cuanto se acaba de decir nos es ya conocido, del tema anterior, ya que el producto escalar en una forma bilineal simétrica definida positiva, no está de más que, por motivos pedagógicos, hagamos las comprobaciones pertinentes.

- La expresión del producto escalar se obtiene del siguiente modo:

$$\bar{x}\cdot\bar{y}=(\sum_i x_i\bar{e}_i)\cdot(\sum_j y_j\bar{e}_j)=\sum_{i,j}(x_i\bar{e}_i)\cdot(y_j\bar{e}_j)=\sum_{i,j}x_iy_j(\bar{e}_i\cdot\bar{e}_j)=\sum_{i,j}x_iy_jg_{ij}$$

  Expresando matricialmente este último resultado, se obtiene directamente que $\bar{x}\cdot\bar{y}=X^tGY$ (nótese que, en rigor, X^tGY es una matriz de tamaño 1×1 cuyo único elemento es el número $\bar{x}\cdot\bar{y}$). Nótese que, como $g_{ij}=\bar{e}_i\cdot\bar{e}_j=\bar{e}_j\cdot\bar{e}_i=g_{ji}$, la matriz G es simétrica.
- Al recurrir a una nueva base y producirse el cambio de coordenadas $X=PX'$ (aquí X y X' son las columnas de coordenadas antiguas y nuevas, respectivamente, del mismo vector $\bar{x}$), resulta que:

$$\bar{x}\cdot\bar{y}=X^tGY=(PX')^tG(PY')=X'^t(P^tGP)Y'$$

  luego P^tGP es la matriz métrica del producto escalar, G', respecto de la nueva base de V.
- Obsérvese finalmente que, como $\bar{x}\cdot\bar{x}>0$ para todo $\bar{x}\neq\bar{o}$, resulta que, la forma bilineal simétrica $(\bar{x},\ \bar{y})\rightarrowtail\bar{x}\cdot\bar{y}$ (es decir, el producto escalar) es definida positiva. Consecuentemente, $G=[g_{ij}]$ es una matriz simétrica definida positiva.

OBSERVACIONES

1. Como el producto escalar es conmutativo, $\bar{x}\cdot\bar{y}=\bar{y}\cdot\bar{x}$, por lo que la matriz métrica G es simétrica, resulta que el producto $\bar{x}\cdot\bar{y}$ puede expresarse de cualquiera de las formas:

$$\bar{x}\cdot\bar{y}=X^tGY\quad;\quad\bar{x}\cdot\bar{y}=Y^tGX$$

2. Sabemos que, en cualquier base de V, la matriz métrica G (simétrica) es definida positiva. Esto, según sabemos, puede expresarse de cualquiera de las siguientes maneras:

   - G es tal que $X^tGX>0$ para toda columna $X\in\mathcal{M}_{n\times 1}$ no nula.
   - $\operatorname{sig}G=(n,\ 0)$; es decir, G es congruente con la matriz unidad, la cual es, pues, la matriz canónica de congruencia de G.

- Se verifica que $\Delta_1 > 0$, $\Delta_2 > 0$, ..., $\Delta_n > 0$, donde Δ_i (para $i = 1, 2, ..., n$) es el determinante

$$\Delta_i = \begin{vmatrix} g_{11} & \cdots & g_{1i} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ g_{i1} & \cdots & g_{ii} \end{vmatrix}$$

(en particular $g_{11} > 0$ y $\det G > 0$, ya que $g_{11} = \Delta_1$ y $\det G = \Delta_n$).

Conviene notar que para que G sea definida positiva es necesario, aun cuando no es suficiente que $g_{11} > 0$, $g_{22} > 0$, ..., $g_{nn} > 0$, puesto que $g_{ii} = \bar{e}_i \cdot \bar{e}_i > 0$.

EJERCICIO

- De un producto escalar entre vectores de $\mathbb{R}^3$, se sabe que, llamando $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ a la base canónica, se verifica que:

$$\bar{e}_1 \cdot \bar{e}_1 = 2 \qquad \bar{e}_2 \cdot \bar{e}_2 = 5 \qquad \bar{e}_3 \cdot \bar{e}_3 = a$$
$$\bar{e}_2 \cdot \bar{e}_3 = -2 \qquad \bar{e}_1 \cdot \bar{e}_3 = -1 \qquad \bar{e}_1 \cdot \bar{e}_2 = 3$$

1. Hallar la expresión de este producto escalar, en la base canónica.
2. Determinar los valores de a para los que lo arriba definido sea, realmente, un producto escalar.
3. Para $a = 3$, hallar la matriz métrica del producto escalar en la base $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ siendo

$$\bar{u}_1 = (1, 0, 0) \quad , \quad u_2 = (1, -2, 0) \quad \text{y} \quad \bar{u}_3 = (1, 2, -3)$$

RESOLUCION

1. La matriz métrica pedida será, por definición, la:

$$G = \begin{bmatrix} \bar{e}_1 \cdot \bar{e}_1 & \bar{e}_1 \cdot \bar{e}_2 & \bar{e}_1 \cdot \bar{e}_3 \\ \bar{e}_2 \cdot \bar{e}_1 & \bar{e}_2 \cdot \bar{e}_2 & \bar{e}_2 \cdot \bar{e}_3 \\ \bar{e}_3 \cdot \bar{e}_1 & \bar{e}_3 \cdot \bar{e}_2 & \bar{e}_3 \cdot \bar{e}_3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & -1 \\ 3 & 5 & -2 \\ -1 & -2 & a \end{bmatrix}$$

Por ello, la expresión del producto escalar será:

$$(x, y, z) \cdot (x', y', z') = [x \quad y \quad z] \begin{bmatrix} 2 & 3 & -1 \\ 3 & 5 & -2 \\ -1 & -2 & a \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} =$$

$$= 2xx' + 5yy' + azz' + 3(xy' + x'y) - (xz' + x'z) - 2(yz' + y'z)$$

2. La condición a imponer es que G sea definida positiva, lo que equivale a que sean positivos los tres determinantes siguientes

$$\Delta_1 = |2| = 2 \qquad \Delta_2 = \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix} = 1 \qquad \Delta_3 = \det G = a - 1$$

Por tanto, lo definido es un producto escalar si y sólo si es $a > 1$.

3. La nueva matriz métrica G' será:

$$G' = P^t G P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -2 & 0 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 3 & -1 \\ 3 & 5 & -2 \\ -1 & -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 2 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -4 & 5 \\ -4 & 10 & -9 \\ 5 & -9 & 31 \end{bmatrix}$$

8.2. NORMAS Y ANGULOS

NORMA DE UN VECTOR

Recuérdese que, cuando se estudian los vectores geométricos, al hablar del «módulo» o «longitud» de un vector $\bar{a}$ y relacionarlo, después, con el producto escalar, se llega a que

$$\bar{a} \cdot \bar{a} = (\text{longitud del vector } \bar{a})^2$$

Vamos a echar mano aquí de esta relación para, apoyándonos en el producto escalar, definir la «norma» de un vector, que es como se llama ahora a lo que, en el caso de los segmentos orientados, era la longitud.

[137]

Sea V un espacio vectorial euclídeo. Para cualquier vector $\bar{x} \in V$, como $\bar{x} \cdot \bar{x} \geq 0$, existe el siguiente número real $\|\bar{x}\|$, que recibe el nombre de *norma* del vector $\bar{x}$:

$$\|\bar{x}\| = \sqrt{\bar{x} \cdot \bar{x}} \quad \text{(definición de norma)}$$

Propiedades de la norma. Si $\bar{x}$ e $\bar{y}$ son vectores del espacio vectorial euclídeo V y para $\lambda \in \mathbb{R}$, se verifica que:

1. $\|\bar{x}\| > 0$ para $\bar{x} \neq \bar{o}$; además $\|\bar{o}\| = 0$.
2. $\|\lambda \bar{x}\| = |\lambda| \, \|\bar{x}\|$.
3. $|\bar{x} \cdot \bar{y}| \leq \|\bar{x}\| \, \|\bar{y}\|$ (desigualdad de Schwarz).
4. $\|\bar{x} + \bar{y}\| \leq \|\bar{x}\| + \|\bar{y}\|$ (desigualdad de Minkowski).

DEMOSTRACION

1. Para $\bar{x} \neq \bar{o}$, sabemos que es $\bar{x} \cdot \bar{x} > 0$, luego $\sqrt{\bar{x} \cdot \bar{x}} > 0$, es decir $\|\bar{x}\| > 0$. Nótese que $\|\bar{o}\| = \sqrt{\bar{o} \cdot \bar{o}} = \sqrt{0} = 0$.
2. $\|\lambda\bar{x}\| = [(\lambda\bar{x}) \cdot (\lambda\bar{x})]^{1/2} = [\lambda^2(\bar{x} \cdot \bar{x})]^{1/2} = \sqrt{\lambda^2}\sqrt{\bar{x} \cdot \bar{x}} = |\lambda|\,\|\bar{x}\|$.
3. Si $\bar{x} = \bar{o}$ o si $y = \bar{o}$, esta propiedad es evidente, pues dice que es $0 \leqslant 0$; supondremos, pues, que es $\bar{x} \neq \bar{o}$ e $\bar{y} \neq \bar{o}$. Recurramos a uno cualquiera de los dos vectores

$$\bar{v} = \frac{1}{\|\bar{x}\|}\bar{x} \pm \frac{1}{\|\bar{y}\|}\bar{y}$$

y expresemos que se verifica que $\bar{v} \cdot \bar{v} \geqslant 0$; haciendo esto, obtenemos que:

$$0 \leqslant \left(\frac{\bar{x}}{\|\bar{x}\|} \pm \frac{\bar{y}}{\|\bar{y}\|}\right) \cdot \left(\frac{\bar{x}}{\|\bar{x}\|} \pm \frac{\bar{y}}{\|\bar{y}\|}\right) = \frac{\bar{x} \cdot \bar{x}}{\|\bar{x}\|^2} + \frac{\bar{y} \cdot \bar{y}}{\|\bar{y}\|^2} \pm 2\,\frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\|\,\|\bar{y}\|} =$$

$$= 2 \pm 2\,\frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\|\,\|\bar{y}\|} \quad , \quad \text{luego } \pm\bar{x} \cdot \bar{y} \leqslant \|x\|\,\|y\|$$

y esta última relación es equivalente a la que queríamos probar.
4. Echando mano de la desigualdad de Schwarz, podemos poner:

$$\|\bar{x} + \bar{y}\|^2 = (\bar{x} + \bar{y}) \cdot (\bar{x} + \bar{y}) = \|x\|^2 + 2(\bar{x} \cdot \bar{y}) + \|y\|^2 \leqslant$$
$$\leqslant \|\bar{x}\|^2 + 2\|\bar{x}\|\,\|\bar{y}\| + \|\bar{y}\|^2 = [\|\bar{x}\| + \|\bar{y}\|]^2$$

y, extrayendo la raíz cuadrada, se obtiene la desigualdad de Minkowski.

CONSECUENCIAS

1. Se dice que un vector $\bar{u}$ es *unitario* o normalizado si $\|\bar{u}\| = 1$. Si $\bar{x}$ es un vector no nulo, entonces los siguientes vectores $\bar{u}_1$ y $\bar{u}_2$, que tienen la dirección de $\bar{x}$, son unitarios:

$$\bar{u}_1 = \frac{\bar{x}}{\|\bar{x}\|} \quad \text{y} \quad \bar{u}_2 = -\frac{\bar{x}}{\|\bar{x}\|}$$

En efecto: las normas de estos vectores valen 1, ya que:

$$\|\bar{u}_i\|^2 = \bar{u}_i \cdot \bar{u}_i = \left(\pm\frac{\bar{x}}{\|\bar{x}\|}\right) \cdot \left(\pm\frac{\bar{x}}{\|\bar{x}\|}\right) = \frac{\bar{x} \cdot \bar{x}}{\|\bar{x}\|^2} = \frac{\|\bar{x}\|^2}{\|\bar{x}\|^2} = 1$$

2. Para vectores $\bar{x}$ e $\bar{y}$ de un espacio vectorial euclídeo, se verifica que:

$$[\bar{x} \text{ e } \bar{y} \text{ son linealmente dependientes}] \Leftrightarrow |\bar{x} \cdot \bar{y}| = \|x\|\,\|y\|$$

Así ocurre: Si $\bar{x}$ o $\bar{y}$ son nulos la propiedad es evidente. Nótese que dos vectores son linealmente dependientes si y sólo si $\|\bar{y}\|\bar{x} = \pm\|\bar{x}\|\bar{y}$, esto es, si y sólo si $\|\bar{y}\|\bar{x} \mp \|\bar{x}\|\bar{y} = \bar{o}$, esto es. si y sólo si la norma del vector $\|\bar{y}\|\bar{x} \mp \|\bar{x}\|\bar{y}$ es nula, o sea, si y sólo si:

$$0 = (\|\bar{y}\|\bar{x} \mp \|\bar{x}\|\bar{y}) \cdot (\|\bar{y}\|\bar{x} \mp \|\bar{x}\|\bar{y}) = 2\|\bar{x}\|^2\|\bar{y}\|^2 \mp 2\|\bar{x}\|\,\|\bar{y}\|(\bar{x} \cdot \bar{y})$$

lo que equivale a $\|\bar{x}\|\,\|\bar{y}\| = \pm(\bar{x} \cdot \bar{y})$, como había que comprobar.

3. Para vectores $\bar{x}$ e $\bar{y}$ de un espacio vectorial euclídeo, se verifica que:

$$\|\bar{x} - \bar{y}\| \geq \big|\|\bar{x}\| - \|\bar{y}\|\big|$$

En efecto: de acuerdo con la propiedad triangular, es:

$$\left.\begin{array}{ll} \|\bar{x}\| = \|(x-y)+y\| \leq \|x-y\| + \|y\| & \text{luego } \|\bar{x}\| - \|\bar{y}\| \leq \|\bar{x} - \bar{y}\| \\ \|\bar{y}\| = \|(y-x)+x\| \leq \|x-y\| + \|x\| & \text{luego } \|\bar{y}\| - \|\bar{x}\| \leq \|\bar{x} - \bar{y}\| \end{array}\right\}$$

las dos últimas relaciones equivalen a la que había que comprobar.

EJEMPLOS

1. En $\mathbb{R}^n$, con el producto escalar canónico, la norma de un vector vale:

$$\|(x_1, x_2, \ldots, x_n)\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

En este caso, la desigualdad de Schwartz permite escribir:

$$(x_1y_1 + x_2y_2 + \cdots + x_ny_n)^2 \leq (x_1^2 + x_2^2 + \cdots + x_n^2)(y_1^2 + y_2^2 + \cdots + y_n^2)$$

2. En el espacio vectorial de las funciones reales que son continuas un intervalo compacto $[a, b]$, con el producto escalar

$$(f \mid g) = \int_a^b f(x)g(x)\,dx$$

la norma de una función f vale:

$$\|f\| = \left[\int_a^b f(x)^2\,dx\right]^{1/2}$$

Aquí, la desigualdad de Schwarz dice que

$$\left(\int_a^b f(x)g(x)\,dx\right)^2 \leq \left(\int_a^b f(x)^2\,dx\right)\left(\int_a^b g(x)^2\,dx\right)$$

3. En el espacio vectorial $\mathcal{M}_{n\times n}$, de las matrices de tamaño $n \times n$, con el producto escalar $\langle A, B\rangle = \text{traza}(A\underline{B}^t)$, la norma de una matriz $A = [a_{ij}]$ vale

$$\|A\| = \sum_{i,j=1}^{n} a_{ij}^2$$

La desigualdad de Schwartz es, en este ejemplo:

$$\left(\sum_{i,j=1}^{n} a_{ij}b_{ij}\right)^2 \leqslant \left(\sum_{i,j=1}^{n} a_{ij}^2\right)\left(\sum_{i,j=1}^{n} b_{ij}^2\right)$$

OBSERVACION (ESPACIO NORMADO)

La norma se puede definir, sin necesidad de recurrir a un producto escalar, generalizando lo aquí dicho. Para ello, se toman, como punto de partida, aquellas propiedades, de las antes demostradas, que son propias de la norma, es decir, en las que no interviene el producto escalar. En concreto, se define:

Se llama *espacio vectorial normado* a un espacio vectorial real V en el que hay definida una *norma*, entendiendo como tal a toda aplicación

$$\| \ \| : V \to \mathbb{R}, \quad \bar{x} \mapsto \|\bar{x}\|$$

para la que se verifica que, para cualesquiera $\bar{x}, \bar{y} \in V$ y $\lambda \in \mathbb{R}$, es:

1. $\|\bar{x}\| > 0$ para $\bar{x} \neq \bar{o}$; además, $\|\bar{o}\| = 0$.
2. $\|\lambda\bar{x}\| = |\lambda|\,\|\bar{x}\|$.
3. $\|\bar{x} + \bar{y}\| \leqslant \|\bar{x}\| + \|\bar{y}\|$.

Evidentemente, todo espacio vectorial euclídeo es un espacio vectorial normado. El recíproco no se verifica en general.

EJERCICIO

Compruébese que, si $\bar{x}, \bar{y}, \bar{z}$ son vectores de un espacio vectorial euclídeo, se verifica que:

a) $\|\bar{x} + \bar{y} + \bar{z}\|^2 = \|\bar{x}\|^2 + \|\bar{y}\|^2 + \|\bar{z}\|^2 + 2\bar{x}\cdot\bar{y} + 2\bar{x}\cdot\bar{z} + 2\bar{y}\cdot\bar{z}$.

b) $\|\bar{x} + \bar{y}\|^2 + \|x - y\|^2 = 2\|\bar{x}\|^2 + 2\|\bar{y}\|^2$.

COMPROBACION

a) $\|\bar{x} + \bar{y} + \bar{z}\|^2 = (\bar{x} + \bar{y} + \bar{z})\cdot(\bar{x} + \bar{y} + \bar{z}) = \bar{x}\cdot\bar{x} + \bar{y}\cdot\bar{y} + \bar{z}\cdot\bar{z} +$
$+ (\bar{x}\cdot\bar{y} + \bar{y}\cdot\bar{x}) + (\bar{x}\cdot\bar{z} + \bar{z}\cdot\bar{x}) + (\bar{y}\cdot\bar{z} + \bar{z}\cdot\bar{y}) = \|\bar{x}\|^2 + \|\bar{y}\|^2 +$
$+ \|\bar{z}\|^2 + 2\bar{x}\cdot\bar{y} + 2\bar{x}\cdot\bar{z} + 2\bar{y}\cdot\bar{z}$.

b) $\|\bar{x}+\bar{y}\|^2+\|x-y\|^2=(\bar{x}+\bar{y})\cdot(\bar{x}+\bar{y})+(\bar{x}-\bar{y})\cdot(\bar{x}-\bar{y})=$

$=[\|\bar{x}\|^2+\|\bar{y}\|^2+2\bar{x}\cdot\bar{y}]+[\|\bar{x}\|^2+\|\bar{y}\|^2-2\bar{x}\cdot\bar{y}]=2\|\bar{x}\|^2+2\|\bar{y}\|^2$

EJERCICIO

En un espacio vectorial euclídeo V, se llama *distancia* del vector $\bar{x}$ al vector $\bar{y}$ al siguiente número real $d(\bar{x}, \bar{y})$:

$$d(\bar{x}, \bar{y})=\|\bar{x}-\bar{y}\|$$

Demuéstrese que, para cualesquiera vectores $\bar{x}, \bar{y}, \bar{z}\in V$, se verifican las siguientes propiedades:

1. $d(\bar{x}, \bar{y})>0$ si $\bar{x}\neq\bar{y}$; además, $d(\bar{x}, \bar{x})=0$.
2. $d(\bar{x}, \bar{y})=d(\bar{y}, \bar{x})$.
3. $d(\bar{x}, \bar{z})\leq d(\bar{x}, \bar{y})+d(\bar{y}, \bar{z})$ (desigualdad triangular).

RESOLUCION

1. Si $\bar{x}\neq\bar{y}$, entonces $\bar{u}=\bar{x}-\bar{y}\neq\bar{o}$, con lo que

$$d(\bar{x}, \bar{y})=\|\bar{x}-\bar{y}\|=\|\bar{u}\|>0$$

Además

$$d(\bar{x}, \bar{x})=\|\bar{x}-\bar{x}\|=\|\bar{o}\|=0$$

2. $d(\bar{x}, \bar{y})=\|\bar{x}-\bar{y}\|=\|-(\bar{y}-\bar{x})\|=|-1|\,\|\bar{y}-\bar{x}\|=\|\bar{y}-\bar{x}\|=d(\bar{y}, \bar{x})$.
3. $d(\bar{x}, \bar{z})=\|\bar{x}-\bar{z}\|=\|(\bar{x}-\bar{y})+(\bar{y}-\bar{z})\|\leq\|\bar{x}-\bar{y}\|+\|\bar{y}-\bar{z}\|=$
$=d(\bar{x}, \bar{y})+d(\bar{y}, \bar{z})$.

COSENO DEL ANGULO DE DOS VECTORES

El concepto de ángulo, definido con exactitud, y el análisis detallado de cuanto ello lleva consigo se ha sacado de aquí (véase Apéndice 11) pues, además de precisar de conceptos que aún no hemos tratado, se considera que son cuestiones que, por su dificultad, nos desbordan en un momento, como éste, en el que se comienza a estudiar todo ello. Para empezar a tratar de estas cuestiones, nos va a ser suficiente con los conocimientos que, sobre ángulos, se suponen conocidos de estudios anteriores.

[138]

Sean $\bar{x}$ e $\bar{y}$ dos vectores no nulos de un espacio vectorial euclídeo V. El ángulo que forma el vector $\bar{x} \neq \bar{o}$ con el vector $\bar{y} \neq \bar{o}$ (*), que se denota poniendo áng $(\bar{x}, \bar{y})$ ó $\widehat{(\bar{x}, \bar{y})}$, queda caracterizado por su coseno, que vale:

$$\cos(\bar{x}, \bar{y}) = \frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\| \, \|\bar{y}\|}$$

Obviamente, recurriendo al ángulo de $\bar{x} \neq \bar{o}$ con $\bar{y} \neq \bar{o}$, su producto escalar se puede expresar en la forma clásica:

$$\bar{x} \cdot \bar{y} = \|\bar{x}\| \, \|\bar{y}\| \cos(\bar{x}, \bar{y})$$

La medida del ángulo áng $(\bar{x}, \bar{y})$ es de la forma $\theta + 2k\pi$ (donde $k \in \mathbb{Z}$) para un cierto $\theta \in [0, \pi]$.

Si $\dim V = n$ y $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una base de V, entonces los elementos de la matriz métrica $G = [g_{ij}]$ del producto escalar en dicha base (véase [136]) se pueden poner en la forma

$$g_{ij} = \|\bar{e}_i\| \, \|\bar{e}_j\| \cos(\bar{e}_i, \bar{e}_j)$$

Se llaman *cosenos directores* del vector $\bar{x} \neq \bar{o}$ en la anterior base a los cosenos de los ángulos áng $(\bar{x}, \bar{e}_i)$ para $i = 1, 2, ..., n$.

(*) La definición precisa de ángulo se puede estudiar, conociendo antes algo acerca de las rotaciones en el plano, en el Apéndice 11.

COMPROBACION

Nótese que, como se supone $\bar{x} \neq \bar{o}$ e $\bar{y} \neq \bar{o}$, la desigualdad de Schwartz se puede escribir en la forma

$$\frac{|\bar{x} \cdot \bar{y}|}{\|\bar{x}\| \, \|\bar{y}\|} \leq 1 \qquad \text{o sea} \qquad -1 \leq \frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\| \, \|\bar{y}\|} \leq 1$$

Por tanto, siempre existe un y sólo un $\theta \in [0, \pi]$ tal que

$$\cos\theta = \frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\| \, \|\bar{y}\|}$$

Resulta, por tanto, que la expresión dada en el enunciado para el $\cos(\bar{x}, \bar{y})$ tiene sentido.

OBSERVACIONES

1. La expresión que da $\cos(\bar{x}, \bar{y})$ puede interpretarse como una manera de determinar cuál es el ángulo que forma $\bar{x}$ con $\bar{y}$. Por ello, se suele decir que el ángulo $\text{áng}(\bar{x}, \bar{y})$ es aquel cuyo coseno es $(\bar{x} \cdot \bar{y}) : \|\bar{x}\| \, \|\bar{y}\|$.
2. Dos vectores $\bar{x} \neq \bar{o}$ e $\bar{y} \neq \bar{o}$ son linealmente dependientes si y sólo si el ángulo $\text{áng}(\bar{x}, \bar{y})$ mide 0 o π. En efecto: recuérdese que $\bar{x}$ e $\bar{y}$ son linealmente dependientes si y sólo si $|\bar{x} \cdot \bar{y}| = \|x\| \, \|\bar{y}\|$ (véase la segunda consecuencia de [137]), esto es, si y sólo si $\cos(\bar{x}, \bar{y}) = \pm 1$, lo que equivale a decir que el $\text{áng}(\bar{x}, \bar{y})$ mide 0 o π.
3. Los ángulos $\text{áng}(\bar{x}, \bar{y})$ y $\text{áng}(\lambda\bar{x}, \bar{y})$, para $\bar{x} \neq \bar{o}$, $\bar{y} \neq \bar{o}$ y $\lambda \neq 0$, son iguales si es $\lambda > 0$ y suplementarios si es $\lambda < 0$. En efecto:

$$\cos(\lambda\bar{x}, \bar{y}) = \frac{(\lambda\bar{x}) \cdot \bar{y}}{\|\lambda\bar{x}\| \, \|\bar{y}\|} = \frac{\lambda}{|\lambda|} \, \frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\| \, \|\bar{y}\|} = \frac{\lambda}{|\lambda|} \cos(\bar{x}, \bar{y})$$

   Luego para $\lambda > 0$ ambos ángulos son iguales y para $\lambda < 0$ son suplementarios (sus cosenos son opuestos).
4. Sea V un espacio vectorial euclídeo de dimensión n, sea $(\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n)$ una base de V y sea $G = [g_{ij}]$ la matriz métrica del producto escalar en dicha base. Para cualquiera que sea el vector $\bar{x} \in V$, además de las coordenadas $(x_1, x_2, \ldots, x_n)$ de $\bar{x}$ que hemos venido manejando, se definen unas nuevas coordenadas $(x_1^*, x_2^*, \ldots, x_n^*)$ mediante $x_i^* = \bar{x} \cdot \bar{e}_i$. A las primeras se las llama contravariantes y a las nuevas covariantes:

$$\left.\begin{array}{l}(x_1, x_2, \ldots, x_n) \text{ sistema de coordenadas} \\ \textit{contravariantes} \text{ o paralelas del vector } \bar{x}\end{array}\right\} \Leftrightarrow \bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n$$

$$\left.\begin{array}{l}(x_1^*, x_2^*, \ldots, x_n^*) \text{ sistema de coordenadas} \\ \textit{covariantes} \text{ o perpendiculares del vector } \bar{x}\end{array}\right\} \Leftrightarrow x_i^* = \bar{x} \cdot \bar{e}_i \; (i = 1, 2, \ldots, n)$$

   De acuerdo con la definición anterior, se puede poner (para $j = 1, 2, \ldots, n$):

$$x_j^* = \bar{x} \cdot \bar{e}_j = (\sum_i x_i\bar{e}_i) \cdot \bar{e}_j = \sum_i (\bar{e}_i \cdot \bar{e}_j)x_i = \sum_i g_{ij}x_i$$

   que, recurriendo a las matrices columna X y X^* de las coordenadas contravariantes y covariantes, conduce a:

$$X^* = GX \quad \text{ó} \quad X = G^{-1}X^*$$

   Así que se pueden calcular unas coordenadas en función de las otras y cualquiera de los dos sistemas de coordenadas permite determinar inequívocamente al vector.

EJERCICIO

Sea V un espacio vectorial euclídeo tridimensional y sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ una base de V de lo que se sabe que:

$$\|\bar{e}_1\| = 1 \quad , \quad \|\bar{e}_2\| = \sqrt{2} \quad , \quad \|\bar{e}_3\| = 2$$

$$(\widehat{e_2, e_3}) = 90° \quad , \quad (\widehat{e_1, e_3}) = 60° \quad , \quad (\widehat{\bar{e}_1, \bar{e}_2}) = 135°$$

1. Hallar la matriz métrica del producto escalar en la base dada.
2. Hallar el ángulo que forman los vectores $\bar{x}$ e $\bar{y}$ siguientes:

$$\bar{x} = \bar{e}_1 - \bar{e}_2 + \bar{e}_3 \quad , \quad \bar{y} = 2\bar{e}_1 - \bar{e}_3$$

RESOLUCION

1. Los elementos de la matriz métrica $G = [g_{ij}]$ son:

$$g_{11} = \bar{e}_1 \cdot \bar{e}_1 = \|\bar{e}_1\|^2 = 1 \quad g_{23} = \bar{e}_2 \cdot \bar{e}_3 = \|\bar{e}_2\| \, \|\bar{e}_3\| \cos(\bar{e}_2, \bar{e}_3) = 0$$
$$g_{22} = \bar{e}_2 \cdot \bar{e}_2 = \|\bar{e}_2\|^2 = 2 \quad g_{13} = \bar{e}_1 \cdot \bar{e}_3 = \|\bar{e}_1\| \, \|\bar{e}_3\| \cos(\bar{e}_1, \bar{e}_3) = 1$$
$$g_{33} = \bar{e}_3 \cdot \bar{e}_3 = \|\bar{e}_3\|^2 = 4 \quad g_{12} = \bar{e}_1 \cdot \bar{e}_2 = \|\bar{e}_1\| \, \|\bar{e}_2\| \cos(\bar{e}_1, \bar{e}_2) = -1$$

Por tanto, G valdrá

$$G = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 4 \end{bmatrix}$$

2. Calculando primero $\|\bar{x}\|$, $\|\bar{y}\|$ y $\bar{x} \cdot \bar{y}$, se obtiene (X e Y son las columnas de coordenadas de $\bar{x}$ e $\bar{y}$):

$$\bar{x} \cdot \bar{x} = X^t G X = 11 \quad , \quad \|\bar{x}\| = \sqrt{11}$$
$$\bar{y} \cdot \bar{y} = Y^t G Y = 4 \quad , \quad \|\bar{y}\| = 2$$
$$\bar{x} \cdot \bar{y} = X^t G Y = 1$$

$$\cos(\bar{x}, \bar{y}) = \frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\| \, \|\bar{y}\|} = \frac{1}{2\sqrt{11}} \simeq 0{,}151 \quad , \quad (\widehat{\bar{x}, \bar{y}}) \simeq 81{,}33°$$

β VECTORES ORTOGONALES Y ORTONORMALES

Según bien sabemos, el producto escalar es la herramienta indispensable para poder manejar las cuestiones euclídeas, longitudes y ángulos; de ahí que sea muy deseable conseguir que su expresión analítica, en coordenadas, sea lo más simple posible. Dicho de otro modo, interesa que la matriz métrica sea cuanto

más sencilla mejor. Pues bien, como el producto escalar es una forma bilineal simétrica, existen bases en las que la matriz métrica es diagonal; estas bases se llaman ortogonales. Pero, es más, como el producto escalar es, además, definido positivo, hay bases en las que dicha matriz es la unidad; estas últimas bases se llaman ortonormales.

Por otra parte la ortogonalidad, o perpendicularidad, entre vector no es un asunto nuevo; recuérdese que dos vectores geométricos $\bar{a} \neq \bar{o}$ y $\bar{b} \neq \bar{o}$ se decían ortogonales si el ángulo que formaban era recto, es decir, si $\cos(\bar{a}, \bar{b}) = 0$; esto, a su vez, equivale a decir que $\bar{a} \cdot \bar{b} = 0$. Aquí tomaremos esta relación para definir la perpendicularidad (obsérvese que, con ello, se considera que el vector nulo es ortogonal a cualquier vector), con lo que, nuevamente, estamos generalizando, a espacios vectoriales en general, lo que conocíamos del caso concreto de los «segmentos orientados».

8.3. BASES ORTONORMALES

SISTEMAS ORTOGONALES Y ORTONORMALES DE VECTORES

[139]

Sea V un espacio vectorial euclídeo. Se dice que $\bar{x} \in V$ e $\bar{y} \in V$ son *vectores ortogonales* si $\bar{x} \cdot \bar{y} = 0$; cuando así ocurra, se pondrá $\bar{x} \perp \bar{y}$. Se dice que $\bar{x} \in V$ es un *vector normalizado* o *unitario* si $\|\bar{x}\| = 1$.

Dado un sistema $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$, de vectores de V, se dice que:

$$\left.\begin{array}{l}(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p) \text{ es} \\ \textit{sistema ortogonal}\end{array}\right\} \text{ si } \bar{u}_i \cdot \bar{u}_j = 0 \text{ para } i \neq j$$

para $(i, j = 1, 2, ..., n)$

$$\left.\begin{array}{l}(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p) \text{ es} \\ \textit{sistema ortonormal}\end{array}\right\} \text{ si } \bar{u}_i \cdot \bar{u}_j = \begin{cases} 0 \text{ para } i \neq j \\ 1 \text{ para } i = j \end{cases}$$

Se verifica que, si $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es un sistema ortogonal de vectores no nulos de V o, en particular, si es un sistema ortonormal, entonces este sistema $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es linealmente independiente.

DEMOSTRACION

Obsérvese primeramente que todo vector de un sistema ortonormal es unitario y, por tanto, no nulo. Así, pues, la propiedad sólo es necesario demostrarla en el supuesto de ser $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ ortogonal y de vectores no nulos.

Si para unos ciertos escalares $\lambda_1, \lambda_2, ..., \lambda_p \in \mathbb{R}$ se verifica que:

$$\lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p = \bar{o}$$

multiplicando escalarmente por $\bar{u}_i$ (para $i = 1, 2, ..., n$), como $\bar{u}_j \cdot \bar{u}_i = 0$ para $i \neq j$, nos encontramos con que $\lambda_i(\bar{u}_i \cdot \bar{u}_i) = \bar{u}_i \cdot \bar{o} = 0$, ahora bien, por ser $\bar{u}_i \neq \bar{o}$, se verifica que $\bar{u}_i \cdot \bar{u}_i \neq 0$ lo que, junto a la anterior igualdad, conduce a $\lambda_i = 0$ (para $i = 1, 2, ..., n$), que prueba la independencia lineal de $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$.

OBSERVACIONES

- Dos vectores no nulos $\bar{x} \neq \bar{o}$ e $\bar{y} \neq \bar{o}$ son ortogonales si y sólo si forman ángulo recto, pues en este caso $\bar{x} \cdot \bar{y} = 0$ equivale a $\cos(\bar{x}, \bar{y}) = 0$. El vector nulo es ortogonal a cualquier vector $\bar{x}$, ya que $\bar{o} \cdot \bar{x} = 0$. Nótese también que el vector nulo es el único vector que es ortogonal a todos los vectores del espacio.
- Supongamos que $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es un sistema ortogonal, es decir, tal que dos cualesquiera de sus vectores son ortogonales. Si, además, todos sus vectores son no nulos, el sistema se puede «normalizar», lo que significa que los vectores

$$\bar{v}_1 = \frac{\bar{u}_1}{\|\bar{u}_1\|}, \quad \bar{v}_2 = \frac{\bar{u}_2}{\|\bar{u}_2\|}, \quad ..., \quad \bar{v}_p = \frac{\bar{u}_p}{\|\bar{u}_p\|}$$

  forman un sistema ortonormal. Nótese que cada uno de los vectores $\bar{v}_i$, que es unitario, tiene la misma dirección y el mismo sentido que el correspondiente vector $\bar{u}_i$.
- Como el producto escalar es una forma bilineal simétrica, resulta que los vectores $\bar{x}, \bar{y} \in V$ son ortogonales si y sólo si son conjugados respecto de la forma bilineal simétrica $f: V \times V \to \mathbb{R}$, $f(\bar{x}, \bar{y}) = \bar{x} \cdot \bar{y}$.

EJERCICIO (TEOREMA DE PITAGORAS)

En un espacio vectorial euclídeo V, dos vectores $\bar{x}, \bar{y} \in V$ son ortogonales si, y sólo si, para ellos se verifica la relación de Pitágoras, esto es:

$$\bar{x} \perp \bar{y} \Leftrightarrow \|\bar{x} + \bar{y}\|^2 = \|\bar{x}\|^2 + \|\bar{y}\|^2$$

RESOLUCION

Para cualesquiera vectores $\bar{x}, \bar{y} \in V$, se verifica que:

$$\|\bar{x}+\bar{y}\|^2 = (\bar{x}+\bar{y})\cdot(\bar{x}+\bar{y}) = \bar{x}\cdot\bar{x} + 2\bar{x}\cdot\bar{y} + \bar{y}\cdot\bar{y} = \|\bar{x}\|^2 + 2\bar{x}\cdot\bar{y} + \|\bar{y}\|^2$$

Por consiguiente, se verifican las siguientes equivalencias:

$$\|\bar{x}+\bar{y}\|^2 = \|\bar{x}\|^2 + \|\bar{y}\|^2 \quad \Leftrightarrow \quad \bar{x}\cdot\bar{y} = 0 \quad \Leftrightarrow \quad \bar{x} \perp \bar{y}$$

EJEMPLO

Sea V el espacio vectorial de las funciones reales que están definidas y son continuas en el intervalo $[-\pi, \pi]$. Considérese, en V, el producto escalar definido mediante:

$$(f \mid g) = \int_{-\pi}^{\pi} f(x)g(x)\,dx$$

El siguiente sistema de funciones es un sistema ortogonal:

$$\{1, \operatorname{sen} x, \cos x, \operatorname{sen} 2x, \cos 2x, \ldots, \operatorname{sen} px, \cos px\}$$

Así ocurre debido a que son nulos todos los productos escalares siguientes:

$$(\operatorname{sen} hx \mid \operatorname{sen} kx) \quad \text{y} \quad (\cos hx \mid \cos kx) \quad , \quad \text{para } h \neq k$$

$$(\operatorname{sen} hx \mid \cos kx) \quad , \quad (1 \mid \operatorname{sen} hx) \quad \text{y} \quad (1 \mid \cos kx) \quad , \quad h \text{ y } k \text{ cualesquiera}$$

Nótese que, como las normas de estas funciones son:

$$\|1\| = \sqrt{2\pi} \quad , \quad \|\operatorname{sen} hx\| = \sqrt{\pi} \quad , \quad \|\cos hx\| = \sqrt{\pi}$$

el siguiente sistema de funciones es ortonormal:

$$\left\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\operatorname{sen} hx, \frac{1}{\sqrt{\pi}}\cos hx \;/\; h = 1, 2, \ldots, p\right\}$$

BASES ORTONORMALES

[140]

Sea V un espacio vectorial euclídeo de dimensión finita. A una base de V que sea sistema ortogonal u ortonormal se la llamará, respectivamente, *base ortogonal* o *base ortonormal* de V. Se verifica que:

1. Una base de V es una base ortogonal o una base ortonormal si, y sólo si, en ella la matriz métrica del producto escalar es, respectivamente, una matriz diagonal o la matriz unidad.
2. Todo espacio vectorial euclídeo de dimensión finita V tiene bases ortonormales.
3. Para hallar una base ortonormal $B_0 \equiv (\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V, a partir de una base cualquiera $B = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$, bastará con(*) diagonalizar por congruencia la matriz métrica G (correspondiente a la base dada B), obteniendo que $I = P^t G P$; las columnas de P son las columnas de coordenadas de los vectores de B_0 respecto de la base B. Se puede conseguir que P sea triangular superior, con lo que $\mathcal{V}(\bar{e}_1, ..., \bar{e}_p) = \mathcal{V}(\bar{u}_1, ..., \bar{u}_p)$ para $p = 1, 2, ..., n$.

(*) No es este el único «método de ortonormalización de una base». Otro método, equivalente a éste es el de Gram-Schmidt, del que se habla un poco más adelante (véase [141]).

DEMOSTRACION

1. Recuérdese que, en una base $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ de V, la matriz métrica del producto escalar es la $G = [g_{ij}]$ con $g_{ij} = \bar{u}_i \cdot \bar{u}_j$ (para $i, j = 1, 2, ..., n$). Esta matriz es diagonal si $\bar{u}_i \cdot \bar{u}_j = 0$ para $i \neq j$, es decir, si la anterior base es ortogonal. La matriz G es la unidad si $\bar{u}_i \cdot \bar{u}_j = 0$ para $i \neq j$ y $\bar{u}_i \cdot \bar{u}_j = 1$ para $i = j$, es decir, si la base es ortonormal.
2. Para comprobar que existen bases ortonormales, nada mejor que obtener alguna de ellas. A eso nos dedicamos en el próximo punto.
3. De acuerdo con la definición de producto escalar y llamando $f(\bar{x}, \bar{y}) = \bar{x} \cdot \bar{y}$, la aplicación $f : V \times V \to \mathbb{R}$ es una forma bilineal simétrica definida positiva; nótese que, como la matriz métrica G es la matriz de f en la base B, entonces G es definida positiva. Por tanto (véase [133]), existe una base B_0 de V en la que la matriz de f es la unidad; esta base es, de acuerdo con lo dicho en el anterior apartado 1, una base ortonormal. No se olvide que, según ya sabemos, se verifica que $I = P^t G P$, para cierta matriz regular P, y que las columnas de P son las columnas de coordenadas de los vectores de la base B_0, que estamos buscando, respecto de la base B, de la que partimos.

   La diagonalización $I = P^t G P$, de la matriz G, se puede llevar a cabo mediante operaciones elementales de multitud de formas. Como G es definida positiva y también lo son todas las matrices congruentes con ella,

resulta que los elementos de lugares 11, 22, ..., *nn* de todas estas matrices son no nulos (más exactamente, son positivos). Por ello, en el proceso de diagonalización, mediante operaciones elementales, ninguno de los elementos que sirven de pivote va a ser nulo. Así, pues, las operaciones elementales pueden hacerse todas ellas «de arriba a abajo», es decir, de manera que una línea (fila o columna) nunca se utilice para influir en otra anterior a ella. Procediendo, pues, de este modo, es evidente que se consigue que la matriz P sea triangular superior; además, los elementos de su diagonal son no nulos, ya que es regular. De lo dicho hasta aquí se infiere que el vector $\bar{e}_p$ de la base B_0 (para $p = 1, 2, ..., n$) será de la forma

$$\bar{e}_p = \alpha_1\bar{u}_1 + \alpha_2\bar{u}_2 + \cdots + \alpha_p\bar{u}_p \quad \text{(para ciertas } \alpha_i \in \mathbb{R}\text{, con } \alpha_p \neq 0)$$

lo que permite asegurar que $\mathcal{V}(\bar{e}_1, ... \bar{e}_p) = \mathcal{V}(\bar{u}_1, ..., \bar{u}_p)$.

EJERCICIO

Sean $f: V \to V$ y $f^*: V \to V$ dos endomorfismos, del espacio vectorial euclídeo V. Se dice que f^* es el *endomorfismo adjunto* de f si se verifica que: $f(\bar{x}) \cdot y = \bar{x} \cdot f^*(\bar{y})$, para cualesquiera $\bar{x}, \bar{y} \in V$. Se dice que f es un *endomorfismo simétrico* si se verifica que $f(\bar{x}) \cdot \bar{y} = \bar{x} \cdot f(\bar{y})$, para cualesquiera $\bar{x}, \bar{y} \in V$.

Si V tiene dimensión finita n y A y A^* son las matrices de f y f^* en una base ortonormal de V, pruébese que:

1. f^* es el endomorfismo adjunto de f si, y sólo si, A^* es la matriz traspuesta de A.
2. f es un endomorfismo simétrico si, y sólo si, A es una matriz simétrica ($A^t = A$).

RESOLUCION

Sean X e Y las matrices columna de las coordenadas de los vectores $\bar{x}$ e $\bar{y}$. Las columnas de coordenadas de $f(\bar{x})$, $f^*(\bar{x})$ y $f(\bar{y})$ son:

$$f(\bar{x}): X' = AX \quad , \quad f^*(x): X^* = A^*X \quad , \quad f(\bar{y}): Y' = AY$$

Si tenemos en cuenta que, por ser la base ortonormal, la matriz métrica del producto escalar es la unidad ($G = I$), se tiene:

1. La relación $f(\bar{x}) \cdot \bar{y} = \bar{x} \cdot f^*(\bar{y})$, $\forall \bar{x}, \bar{y} \in V$, equivale a cada una de las:

$$(AX)^t IY = X^t I(A^*Y) \quad , \quad \forall X, Y \in \mathcal{M}_{n\times 1}$$

$$X^t A^t Y = X^t A^* Y \quad , \quad \forall X, Y \in \mathcal{M}_{n\times 1}$$

$$A^t = A^*$$

2. Repitiendo lo anterior para $f^* = f$ y $A^* = A$, se obtiene $A^t = A$.

EJERCICIO

Sea V el espacio vectorial de los polinomios, con una indeterminada x, reales de grado menor que tres. Considérese en V el producto escalar definido mediante:

$$(p(x) \mid q(x)) = \int_{-1}^{1} p(x)q(x)\,dx$$

Hallar una base ortonormal de este espacio vectorial euclídeo.

RESOLUCION

Consideremos la base usual $(1, x, x^2)$ y hallemos la matriz métrica del producto escalar G en esta base:

$$(1 \mid 1) = \int_{-1}^{1} 1^2\,dx = 2 \quad , \quad (1 \mid x) = \int_{-1}^{1} x\,dx = 0 \quad , \quad (1 \mid x^2) = \int_{-1}^{1} x^2\,dx = \frac{2}{3}$$

$$(x \mid x) = \int_{-1}^{1} x^2\,dx = \frac{2}{3} \quad , \quad (x \mid x^2) = \int_{-1}^{1} x^3\,dx = 0 \quad , \quad (x^2 \mid x^2) = \int_{-1}^{1} x^4\,dx = \frac{2}{5}$$

$$G = \begin{bmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{bmatrix}$$

Diagonalicemos por congruencia esta matriz G:

$$[G \mid I] = \left[\begin{array}{ccc|ccc} 2 & 0 & 2/3 & 1 & 0 & 0 \\ 0 & 2/3 & 0 & 0 & 1 & 0 \\ 2/3 & 0 & 2/5 & 0 & 0 & 1 \end{array}\right] \simeq \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2/3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 8/5 & -1 & 0 & 3 \end{array}\right] \simeq$$

$$\simeq \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/\sqrt{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & \sqrt{3/2} & 0 \\ 0 & 0 & 1 & -\sqrt{5/8} & 0 & 3\sqrt{5/8} \end{array}\right] = [I \mid P^t]$$

Por tanto, una base ortonormal de V será la $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ donde

$$e_1 = \frac{1}{\sqrt{2}} \quad , \quad e_2 = \sqrt{\frac{3}{2}}\,x \quad , \quad e_3 = \frac{\sqrt{5}}{2\sqrt{2}}(-1 + 3x^2)$$

Nótese que $(a, bx, c - 3cx^2)$ es una base ortogonal de V para cualesquiera $a \neq 0$, $b \neq 0$ y $c \neq 0$. En particular, lo es $(1, x, -1/2 + 3/2x^2)$; estos polinomios, que valen 1 para $x = 1$, son los tres primeros «polinomios de Legendre», que juegan un importante papel en ecuaciones diferenciales.

OBSERVACION (PROCESO DE ORTONORMALIZACION DE GRAM-SCHMIDT)

[141] Dada una base $B = (\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$, del espacio vectorial euclídeo V de dimensión finita, existen bases ortogonales de V, que denotaremos por $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$, tales que $\mathcal{V}(\bar{e}_1, ..., \bar{e}_p) = \mathcal{V}(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ para $p = 1, 2, ..., n$. Una de estas bases es aquella en la que $\bar{e}_p$ vale:

$$\bar{e}_p = \bar{u}_p - (\alpha_1\bar{e}_1 + \cdots + \alpha_{p-1}\bar{e}_{p-1}) \quad \text{con } \alpha_i = \frac{\bar{u}_p \cdot \bar{e}_i}{\|\bar{e}_i\|^2} \text{ (para } i = 1, ..., p-1)$$

Si se «normaliza» la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$, se obtiene la base ortonormal $(\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n)$, donde $\bar{e}'_p = \bar{e}_p / \|\bar{e}_p\|$; para esta última base también se cumple que $\mathcal{V}(\bar{e}'_1, ..., \bar{e}'_p) = \mathcal{V}(\bar{u}_1, ..., \bar{u}_p)$ para $p = 1, 2, ..., n$.

DEMOSTRACION

Recurramos al método de inducción, sobre p. Para $n = 1$, la propiedad es evidentemente cierta. Suponiendo que la propiedad es cierta para $p - 1$ vectores, veamos que también lo es para p. Admitimos, pues que $(\bar{e}_1, ..., \bar{e}_{p-1})$ es un sistema ortogonal de vectores no nulos tal que $\mathcal{V}(\bar{e}_1, ..., \bar{e}_{p-1}) = \mathcal{V}(\bar{u}_1, ..., \bar{u}_{p-1})$. El vector $\bar{e}_p$ que se define en el enunciado es ortogonal a los $\bar{e}_1, ..., \bar{e}_{p-1}$ ya que, multiplicando escalarmente $\bar{e}_p$ por $\bar{e}_i$ (para $i = 1, ..., p-1$), se obtiene:

$$\bar{e}_p \cdot \bar{e}_i = \bar{u}_p \cdot \bar{e}_i - (\alpha_1\bar{e}_1 + \cdots + \alpha_{p-1}\bar{e}_{p-1}) \cdot \bar{e}_i = \bar{u}_p \cdot \bar{e}_i - \alpha_i(\bar{e}_i \cdot \bar{e}_i) =$$

$$= \bar{u}_p \cdot \bar{e}_i - \frac{\bar{u}_p \cdot \bar{e}_i}{\|\bar{e}_i\|^2}\|\bar{e}_i\|^2 = 0$$

Finalmente, según resulta fácil comprobar con detalle, se verifica que:

$$\mathcal{V}(\bar{e}_1, ..., \bar{e}_{p-1}, \bar{e}_p) = \mathcal{V}(\bar{e}_1, ..., \bar{e}_{p-1}, \bar{u}_p) = \mathcal{V}(\bar{u}_1, ..., \bar{u}_{p-1}, \bar{u}_p)$$

Las afirmaciones que, en el enunciado, se refieren a $(\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n)$ son evidentes.

TEOREMA DE LA BASE ORTONORMAL INCOMPLETA

En un espacio vectorial euclídeo V de dimensión finita, todo sistema ortonormal de vectores puede completarse hasta obtener una base ortonormal. Dicho con mayor precisión:

[142]

Si V es un espacio euclídeo V de dimensión finita n y si $S = (\bar{e}_1, ..., \bar{e}_p)$ es un sistema ortonormal de p vectores de V, donde $p < n$, entonces existe algún sistema $S' = (\bar{e}_{p+1}, ..., \bar{e}_n)$, de $n - p$ vectores de V, tal que $S \cup S'$ es una base ortonormal de V.

DEMOSTRACION

Para empezar, comprobemos que existe un vector $\bar{e}_{p+1} \in V$ tal que $(\bar{e}_1, ..., \bar{e}_p, \bar{e}_{p+1})$ es un sistema ortonormal. Como $p < n$, existe algún vector $\bar{u} \in V$ que es independiente del sistema dado S; consecuentemente, el vector

$$\bar{e} = \bar{u} - (\alpha_1\bar{e}_1 + \alpha_2\bar{e}_2 + \cdots + \alpha_p\bar{e}_p) \quad , \quad \text{donde } \alpha_1, \alpha_2, ..., \alpha_p \in \mathbb{R}$$

es no nulo para cualesquiera que sean estos escalares. Pues bien, existen ciertos $\alpha_1, \alpha_2, ..., \alpha_p \in \mathbb{R}$ para los que $\bar{e}$ es ortogonal a $\bar{e}_1, \bar{e}_2, ..., \bar{e}_p$, es decir, tales que $\bar{e} \cdot \bar{e}_i = 0$ para $i = 1, 2, ..., p$. Así ocurre, ya que esto equivale a:

$$0 = \bar{u} \cdot \bar{e}_i - \sum_j \alpha_j(\bar{e}_j \cdot \bar{e}_i) \quad \text{o sea} \quad \bar{u} \cdot \bar{e}_i = \alpha_i \quad (\text{para } i = 1, 2, ..., p)$$

Para estos valores de los α_i, el vector $\bar{e}$ es, pues ortogonal a los $\bar{e}_1, \bar{e}_2, ..., \bar{e}_p$ y por ello el vector $\bar{e}_{p+1} = \bar{e}/\|\bar{e}\|$ es tal que $(\bar{e}_1, ..., \bar{e}_p, \bar{e}_{p+1})$ es un sistema ortonormal.

Reiterando este proceso $n - p$ veces, nos encontramos con que existe un sistema ortonormal $(\bar{e}_1, ..., \bar{e}_p, \bar{e}_{p+1}, ..., \bar{e}_n)$ que, como es independiente y consta de n vectores, es base (ortonormal) de V, con lo que concluye la demostración.

EJERCICIO

En el espacio vectorial euclídeo canónico $\mathbb{R}^4$, se consideran los siguientes vectores $\bar{e}_1$ y $\bar{e}_2$, que son ortogonales:

$$\bar{e}_1 = (1, 0, -1, 1) \quad \text{y} \quad \bar{e}_2 = (2, 1, 2, 0)$$

Completar el sistema ortogonal $(\bar{e}_1, \bar{e}_2)$ hasta obtener una base ortogonal de $\mathbb{R}^4$.

RESOLUCION

Añadamos dos vectores al sistema $(\bar{e}_1, \bar{e}_2)$ hasta obtener una base (cualquiera) de $\mathbb{R}^4$; tales vectores pueden ser, por ejemplo, los $\bar{u}=(0, 0, 1, 0)$ y $\bar{v}=(0, 0, 0, 1)$. La matriz métrica del producto escalar en la base $(\bar{e}_1, \bar{e}_2, \bar{u}, \bar{v})$ será la:

$$G=\begin{bmatrix} \bar{e}_1\cdot\bar{e}_1 & \bar{e}_1\cdot\bar{e}_2 & \bar{e}_1\cdot\bar{u} & \bar{e}_1\cdot\bar{v} \\ \bar{e}_2\cdot\bar{e}_1 & \bar{e}_2\cdot\bar{e}_2 & \bar{e}_2\cdot\bar{u} & \bar{e}_2\cdot\bar{v} \\ \bar{u}\cdot\bar{e}_1 & \bar{u}\cdot\bar{e}_2 & \bar{u}\cdot\bar{u} & \bar{u}\cdot\bar{v} \\ \bar{v}\cdot\bar{e}_1 & \bar{v}\cdot\bar{e}_2 & \bar{v}\cdot\bar{u} & \bar{v}\cdot\bar{v} \end{bmatrix}=\begin{bmatrix} 3 & 0 & -1 & 1 \\ 0 & 9 & 2 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Diagonalizando (por congruencia) la matriz G haciéndolo de modo que no se alteren los vectores $\bar{e}_1$ y $\bar{e}_2$ (que es posible, pues éstos son ortogonales), se obtiene:

$$[G\,|\,I]\cong\left[\begin{array}{cccc|cccc} 3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 9 & 6 & 0 & 0 & 1 & 0 & 0 \\ 0 & 6 & 6 & 3 & 1 & 0 & 3 & 0 \\ 0 & 0 & 3 & 6 & -1 & 0 & 0 & 3 \end{array}\right]\cong\left[\begin{array}{cccc|cccc} 3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 18 & 9 & 3 & -2 & 9 & 0 \\ 0 & 0 & 9 & 6 & -1 & 0 & 0 & 3 \end{array}\right]\cong$$

$$\cong\left[\begin{array}{cccc|cccc} 3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 18 & 0 & 3 & -2 & 9 & 0 \\ 0 & 0 & 0 & 6 & -5 & 2 & -9 & 6 \end{array}\right]$$

$$\begin{aligned} \bar{e}_1 &= \bar{e}_1 && =(1, 0, -1, 1) \\ \bar{e}_2 &= \bar{e}_2 && =(2, 1, 2, 0) \\ \bar{e}_3 &= 3\bar{e}_1-2\bar{e}_2+9\bar{u} && =(-1, -2, 2, 3) \\ \bar{e}_4 &= -5\bar{e}_1+2\bar{e}_2-9\bar{u}+6\bar{v} && =(-1, 2, 0, 1) \end{aligned}$$

Por tanto, los anteriores vectores $\bar{e}_3$ y $\bar{e}_4$ son solución del problema, ya que $(\bar{e}_1, \bar{e}_2, \bar{e}_3, \bar{e}_4)$ es base ortogonal de $\mathbb{R}^4$.

COORDENADAS Y PRODUCTO ESCALAR EN UNA BASE ORTONORMAL

Como ya se dijo, el interés de las bases ortonormales radica en que, en ellas, la expresión del producto escalar se simplifica al máximo y a que, por tanto, ocurre lo propio con las fórmulas que proporcionan las longitudes y los ángulos. Ha llegado el momento de verificar que las cosas son como se esperaba:

[143]

Sea V un espacio vectorial euclídeo de dimensión finita, en el que se considera una base ortonormal $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$. Llamando $(x_1, x_2, ..., x_n)$ e $(y_1, y_2, ..., y_n)$ a las coordenadas de dos vectores cualesquiera $\bar{x}, \bar{y} \in V$, en la base dada, se verifica que:

1. $\bar{x} \cdot \bar{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$.
2. $\|\bar{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.
3. $x_i = \bar{x} \cdot \bar{e}_i = \|\bar{x}\| \cos(\bar{x}, \bar{e}_i)$ (para $i = 1, 2, ..., n$)

DEMOSTRACION

1. La primera fórmula se obtiene, de la ya conocida $\bar{x} \cdot \bar{y} = \Sigma g_{ij} x_i x_j$ (véase [136]), teniendo en cuenta que ahora la matriz $G = [g_{ij}]$ es la unidad.
2. Como $\|\bar{x}\|^2 = \bar{x} \cdot \bar{x}$, la expresión de $\|\bar{x}\|$ se obtiene de tomar $\bar{x} = \bar{y}$ en la fórmula anterior y extraer luego la raíz cuadrada.
3. Multiplicando escalarmente por $\bar{e}_i$ a los dos miembros de la igualdad

$$\bar{x} = x_1 \bar{e}_1 + x_2 \bar{e}_2 + \cdots + x_n \bar{e}_n$$

como la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es ortonormal, resulta que $\bar{x} \cdot \bar{e}_i = x_i$. Nótese finalmente que $\bar{x} \cdot \bar{e}_i = \|\bar{x}\| \cos(\bar{x}, \bar{e}_i)$.

EJERCICIO

Sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ una base ortonormal de un espacio vectorial euclídeo tridimensional. Compruébese que, para ciertos $\alpha, \beta \in \mathbb{R}$, los tres vectores

$$\bar{u}_1 = \alpha \bar{e}_2 + \beta \bar{e}_3 \quad ; \quad \bar{u}_2 = \alpha \bar{e}_3 + \beta \bar{e}_1 \quad ; \quad \bar{u}_3 = \alpha \bar{e}_1 + \beta \bar{e}_2$$

son unitarios y tales que cada dos de ellos forman un ángulo de 60°.

RESOLUCION

$$\|\bar{u}_1\| = \|\bar{u}_2\| = \|\bar{u}_3\| = \alpha^2 + \beta^2 = 1$$

$$\cos(\bar{u}_1, \bar{u}_2) = \cos(\bar{u}_2, \bar{u}_3) = \cos(\bar{u}_3, \bar{u}_1) = \alpha\beta = 1 \cdot 1 \cdot \cos 60° = 1/2$$

El sistema formado por las ecuaciones $\alpha^2 + \beta^2 = 1$ y $2\alpha\beta = 1$ tiene las soluciones $\alpha = \beta = \pm 1/\sqrt{2}$.

EJERCICIO

Sea $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ una base ortonormal de un espacio vectorial euclídeo V, de dimensión n. Pruébese que, para cualesquiera que sean los vectores $\bar{u}, \bar{v} \in V$, se verifica que

$$\bar{u} \cdot \bar{v} = (\bar{u} \cdot \bar{e}_1)(\bar{v} \cdot \bar{e}_1) + (\bar{u} \cdot \bar{e}_2)(\bar{v} \cdot \bar{e}_2) + \cdots + (\bar{u} \cdot \bar{e}_n)(\bar{v} \cdot \bar{e}_n)$$

RESOLUCION

Sabemos que las coordenadas de u son los escalares $u_i = \bar{u} \cdot \bar{e}_i$ y que las coordenadas de $\bar{v}$ son los escalares $v_i = \bar{v} \cdot \bar{e}_i$ (para $i = 1, 2, ..., n$). Como el producto escalar $\bar{u} \cdot \bar{v}$ vale $\bar{u} \cdot \bar{v} = \sum u_i v_i$, trayendo aquí las anteriores expresiones de u_i y v_i se obtiene la fórmula del enunciado.

8.4. PROYECCION ORTOGONAL

El problema que aquí se plantea es ya conocido de geometría elemental; se trata de lo siguiente: dados un vector $\bar{x}$, de un cierto espacio vectorial euclídeo V, y un subespacio vectorial U, de V (por ejemplo, si $V = \mathbb{R}^3$, U puede ser un plano vectorial), se busca la «proyección ortogonal de $\bar{x}$ sobre U», que es aquel vector $\bar{x}_u$ de U, si existe, tal que $\bar{x} - \bar{x}_u$ es ortogonal a U. En geometría elemental, esta cuestión siempre tenía solución única; ahora, en general, lo sigue teniendo siempre que U tenga dimensión finita.

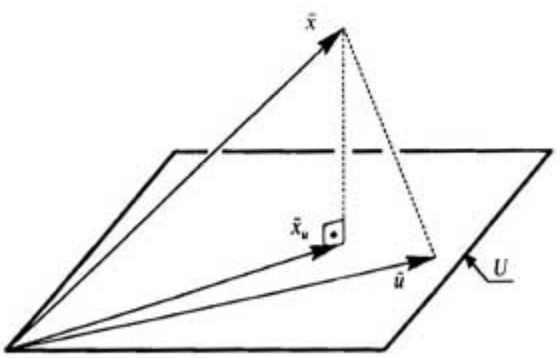


Figura 1.

La proyección ortogonal $\bar{x}_u$ de $\bar{x}$ sobre U se suele llamar también «la mejor aproximación de $\bar{x}$ mediante vectores de U». Ello es debido a que $\|\bar{x} - \bar{x}_u\|$ es la menor de las normas $\|\bar{x} - \bar{u}\|$ para $\bar{u}$ recorriendo U. Esto es, la proyección ortogonal de $\bar{x}$ sobre U va a proporcionar la «mínima distancia» de $\bar{x}$ a U.

SUBESPACIO ORTOGONAL

[144]

Sea V un espacio vectorial euclídeo. Se dice que dos subespacios de V son ortogonales si cualquier vector de uno de ellos es ortogonal a todos los vectores del otro.

Dado un subespacio U, del espacio vectorial euclídeo V, el conjunto

$$U^{\perp} = \{\bar{x} \in V \,/\, \bar{x} \cdot \bar{u} = 0,\ \forall\, \bar{u} \in U\}$$

es el mayor de los subespacios de V que son ortogonales a U y se verifica que $U \cap U^{\perp} = O$. Se dice que $U^{\perp}$ es el *subespacio ortogonal* de U.

Si V tiene dimensión finita, entonces el subespacio ortogonal $U^{\perp}$, de un subespacio cualquiera U de V, es suplementario de U, o sea, $U \oplus U^{\perp} = V$. Para señalar que esto ocurre, se dice que $U^{\perp}$ es el *suplementario ortogonal* de U. Si $(\bar{e}_1, ..., \bar{e}_p)$ es una base ortonormal de U y se la completa hasta obtener una base ortonormal $(\bar{e}_1, ..., \bar{e}_p, ..., \bar{e}_n)$ de V, entonces $U^{\perp} = \mathcal{V}(\bar{e}_{p+1}, ..., \bar{e}_n)$.

DEMOSTRACION

1. El conjunto $U^{\perp}$ es subespacio de V ya que no es vacío (pues $\bar{o}$ le pertenece) y, si $\bar{x}, \bar{y} \in U^{\perp}$ y $\lambda, \mu \in \mathbb{R}$, para cualquiera que sea $\bar{u} \in U$, se verifica que

$$(\lambda\bar{x} + \mu\bar{y}) \cdot \bar{u} = \lambda(\bar{x} \cdot \bar{u}) + \mu(\bar{y} \cdot \bar{u}) = \lambda 0 + \mu 0 = 0$$

luego $\lambda\bar{x} + \mu\bar{y} \in U^{\perp}$, como había que comprobar.

2. La intersección $U \cap U^{\perp}$ es nula ya que, si $\bar{x} \in U$ y $\bar{x} \in U^{\perp}$, deberá ser $\bar{x} \cdot \bar{x} = 0$, o sea, se concluye que $\bar{x} = \bar{o}$.

3. Como U tiene dimensión finita, sabemos (véase [140], 2) que existe una base ortonormal suya $(\bar{e}_1, ..., \bar{e}_p)$. Por el teorema de la base ortonormal incompleta (véase [142]), la base anterior se puede completar hasta obtener una base ortonormal $(\bar{e}_1, ..., \bar{e}_p, ..., \bar{e}_n)$ de V. Los vectores $\bar{e}_{p+1}, ..., \bar{e}_n$ son de $U^{\perp}$, pues son ortogonales a todo vector de U.

   Por otra parte, cualquier vector $\bar{x} \in V$, expresándolo en la anterior base se puede poner:

$$\bar{x} = \bar{u} + \bar{u}_1 \quad , \quad \text{donde} \begin{cases} \bar{u} = x_1\bar{e}_1 + \cdots + x_p\bar{e}_p \in U \\ \bar{u}' = x_{p+1}\bar{e}_{p+1} + \cdots + x_n\bar{e}_n \in U^{\perp} \end{cases}$$

luego $U + U^{\perp} = V$. Como sabemos que $U \cap U^{\perp} = O$, resulta de todo ello que $U \oplus U^{\perp}$.

Veamos, finalmente, que $(\bar{e}_{p+1}, \ldots \bar{e}_n)$ es una base de $U^\perp$. Según sabemos, $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ es un sistema ortonormal de vectores de $U^\perp$ y, por tanto, si no fuese una base de $U^\perp$ habría algún vector no nulo $\bar{e} \in U^\perp$ ortogonal a los $\bar{e}_{p+1}, \ldots, \bar{e}_n$, de donde resultaría que $\bar{e} \in U$, y esto no es posible puesto que $U \cap U^\perp = O$.

EJEMPLO

En el espacio vectorial euclídeo canónico $\mathbb{R}^4$, considérese el subespacio bidimensional

$$U = \{(a, a, b, b) \in \mathbb{R}^4 \,/\, a, b \in \mathbb{R}\}$$

El subespacio suplementario ortogonal de U es el

$$U_1 = \{(\alpha, -\alpha, \beta, -\beta) \in \mathbb{R}^4 \,/\, \alpha, \beta \in \mathbb{R}\}$$

En efecto: Cualquier vector de U_1 es ortogonal a todos los de U, luego $U_1 \subset U^\perp$. Como $\dim U^\perp = 4 - \dim U = 2$ y también $\dim U_1 = 2$, resulta que $U_1 = U^\perp$.

EJERCICIO

Si U es un subespacio de un espacio vectorial V de dimensión finita, compruébese que $(U^\perp)^\perp = U$.

RESOLUCION

Sea $(\bar{e}_1, \ldots, \bar{e}_p)$ una base ortonormal de U, que completada conduce a la base ortonormal $(\bar{e}_1, \ldots, \bar{e}_p, \ldots, \bar{e}_n)$ de V, con lo que $U^\perp = \mathcal{V}(\bar{e}_{p+1}, \ldots, \bar{e}_n)$. Nótese que $(\bar{e}_1, \ldots, \bar{e}_p, \ldots, \bar{e}_n)$ es base de V que completa (ortonormalmente) a la base $(\bar{e}_{p+1}, \ldots, \bar{e}_n)$ de $U^\perp$, de donde resulta que $(\bar{e}_1, \ldots, \bar{e}_p)$ es una base de $(U^\perp)^\perp$; como $(\bar{e}_1, \ldots, \bar{e}_p)$ era base de U, se concluye que $(U^\perp)^\perp = U$.

OBSERVACION

Si V es un espacio vectorial euclídeo que no tiene dimensión finita, un subespacio suyo U y el subespacio ortogonal a este $U^\perp$ pueden no ser suplementarios. Es decir que, aun cuando siempre se verifica que $U \cap U^\perp = O$, la suma $U + U^\perp$ (que es directa) puede no ser todo el espacio V, como ocurre en el siguiente ejemplo:

Sea V el espacio vectorial euclídeo de las funciones polinómicas, definidas en [0, 1], con el siguiente producto escalar:

$$(p \mid q) = \int_0^1 p(x)q(x)\,dx$$

Sea U el subespacio, de V, que forman las funciones polinómicas que tienen nulo el término independiente. No cuesta demasiado trabajo comprobar(*) que no existe ninguna función polinómica no nula que sea ortogonal a todas las de U; dicho de otro modo, se verifica que $U^{\perp} = O$. En este ejemplo es $U + U^{\perp} \neq V$.

EJERCICIO

Sea $f: V \to V$ un endomorfismo, en el espacio vectorial euclídeo V. Supóngase que f es simétrico, esto es, tal que $\bar{x} \cdot f(\bar{y}) = f(\bar{x}) \cdot \bar{y}$ para cualesquiera $\bar{x}$, $\bar{y} \in V$ (véase el primer ejercicio de [140]). Pruébese que el núcleo de f y la imagen de f son dos subespacios suplementarios y ortogonales, es decir, que cada uno de ellos es el suplementario ortogonal del otro.

RESOLUCION

Llamemos $N = \text{Nuc}(f)$ e $I = \text{Im}(f)$. Se verifica que

1. $N \cap I = O$, ya que si $\bar{x} \in N \cap I$, será:
   Como $\bar{x} \in N$, ha de ser $f(\bar{x}) = \bar{o}$; como $\bar{x} \in I$, ha de ser $\bar{x} = f(\bar{y})$ para algún $\bar{y} \in V$; como f es simétrico, es $\bar{x} \cdot f(\bar{y}) = \bar{y} \cdot f(\bar{x})$, esto es:

   $$\bar{x} \cdot \bar{x} = \bar{y} \cdot \bar{o} = 0 \quad , \quad \text{luego } \bar{x} = \bar{o}$$

2. $N \oplus I = V$, ya que $N \cap I = O$ y (véase [082], 4)

   $$\dim N + \dim I = \dim [\text{Nuc}(f)] + \dim [\text{Im}(f)] = \dim V$$

3. N e I son ortogonales ya que, si $\bar{x} \in N$ e $\bar{y} \in I$, se verificará que $f(\bar{x}) = \bar{o}$ e $\bar{y} = f(\bar{z})$ para algún $\bar{z} \in V$ y, por ser f simétrico, es

   $$\bar{x} \cdot \bar{y} = \bar{x} \cdot f(\bar{z}) = \bar{z} \cdot f(\bar{x}) = \bar{z} \cdot \bar{o} = 0$$

   luego $\bar{x}$ e $\bar{y}$ son ortogonales, como había que comprobar.

(*) Dado $p(x) = a_0 + a_1x + \cdots + a_nx^n$, compruébese que $\int_0^1 x^i p(x) = 0$ para $i = 1, 2, \ldots, n+2$ es un sistema lineal, de $n+1$ ecuaciones lineales en las $n+1$ incógnitas $a_0, a_1, \ldots, a_n$, que tiene determinante no nulo y, por ello, su única solución es $a_0 = a_1 = \cdots = a_n = 0$.

PROYECCION ORTOGONAL Y MINIMA DISTANCIA

[145]

Sea V un espacio vectorial euclídeo, en el que se consideran un subespacio $U \subset V$ y un vector $\bar{x} \in V$. Se dice que:

a) El vector $\bar{x}_u$ es *proyección ortogonal* de $\bar{x}$ sobre U si $\bar{x}_u \in U$ y $\bar{x} - \bar{x}_u \in U^\perp$ (el vector $\bar{x} - \bar{x}_u$ es ortogonal a U).

b) El vector $\bar{x}_u \in U$ es el *«más próximo»* a $\bar{x}$ de entre los vectores de U si $\|\bar{x} - \bar{x}_u\|$ es el menor de los valores $\|\bar{x} - \bar{u}\|$ para $\bar{u} \in U$.

Se verifica que un vector $\bar{x}_u$ es la proyección ortogonal de $\bar{x}$ sobre U si, y sólo si, $\bar{x}_u$ es el más próximo a $\bar{x}$ de entre los vectores de U. Esto es:

$$a) \Leftrightarrow b)$$

Dados U y $\bar{x}$, en V, la proyección ortogonal y el vector más próximo son únicos (cuando existen).

DEMOSTRACION

En lo que sigue, nos va a ser útil recurrir a que, para cualesquiera que sean los tres vectores $\bar{x}$, $\bar{u}$ y $\bar{x}_u$, se puede poner

$$\|\bar{x} - \bar{u}\|^2 = \|(\bar{x} - \bar{x}_u) + (\bar{x}_u - \bar{u})\|^2 = [(\bar{x} - \bar{x}_u) + (\bar{x}_u - \bar{u})] \cdot [(\bar{x} - \bar{x}_u) + (\bar{x}_u - \bar{u})] =$$
$$= \|\bar{x} - \bar{x}_u\|^2 + \|\bar{x}_u - \bar{u}\|^2 + 2(\bar{x} - \bar{x}_u) \cdot (\bar{x}_u - \bar{u}) \qquad [I]$$

1. Comprobemos en primer lugar que $a) \Rightarrow b)$. Suponemos, pues, que $\bar{x}_u \in U$ es tal que $\bar{x} - \bar{x}_u \in U^\perp$. Para cualquiera $\bar{u} \in U$, resulta que $\bar{x}_u - \bar{u} \in U$ y, por ello, $\bar{x} - \bar{x}_u$ y $\bar{x}_u - \bar{u}$ son ortogonales, luego su producto escalar es nulo, lo que llevado a la igualdad [I] conduce a:

$$\|\bar{x} - \bar{u}\|^2 = \|\bar{x} - \bar{x}_u\|^2 + \|\bar{x}_u - \bar{u}\|^2 \geq \|\bar{x} - \bar{x}_u\|^2$$

   Así, pues, $\|\bar{x} - \bar{x}_u\|$ ha resultado ser la menor de las $\|\bar{x} - \bar{u}\|$ para $\bar{x} \in U$.

2. Comprobemos ahora que $b) \Rightarrow a)$. Partimos de que $\bar{x}_u \in U$ es tal que $\|\bar{x} - \bar{x}_u\| \leq \|\bar{x} - \bar{u}\|$ para todo $\bar{u} \in U$; por ello, de la igualdad [I] se desprende que:

$$\|\bar{x}_u - \bar{u}\|^2 + 2(\bar{x} - \bar{x}_u) \cdot (\bar{x}_u - \bar{u}) \geq 0 \quad \text{para todo } \bar{u} \in U$$

Nótese que, cuando $\bar{u}$ recorre U, el vector $\bar{u}' = \bar{u} - \bar{x}_u$ también recorre todo U, con lo que la relación anterior se puede expresar de la siguiente forma:

$$\|\bar{u}'\|^2 \geq 2(\bar{x} - \bar{x}_u) \cdot \bar{u} \quad \text{para todo } \bar{u}' \in U \qquad [II]$$

Supongamos que la propiedad a demostrar fuese falsa, esto es, que $\bar{x} - \bar{x}_u$ no fuese ortogonal a algún $\bar{u}_0 \in U$ (obviamente sería $\bar{u}_0 \neq \bar{o}$), o sea, que $(\bar{x} - \bar{x}_u) \cdot \bar{u}_0 \neq 0$ para algún $\bar{u}_0 \in U$; llamemos entonces c al número $c = (\bar{x} - \bar{x}_u) \cdot \bar{u}_0 \neq 0$. Tomemos para vector $\bar{u}'$, de la relación [II], a $\bar{u}' = h\bar{u}_0$ para $h \in \mathbb{R}$ cualquiera; haciendo esto, nos encontramos conque para todo $h \in \mathbb{R}$ habría de ser

$$h^2\|\bar{u}_0\|^2 \geq 2ch \quad , \quad \text{luego } h \geq 2c/\|\bar{u}_0\|^2 > 0$$

que no se verifica para todo $h \in \mathbb{R}$. Esta contradicción nos lleva a que $b) \Rightarrow a)$ es verdadera.

3. Para concluir, veamos que no hay más de un $\bar{x}_n$. Nótese que, como $a) \Leftrightarrow b)$, bastará con probar que la proyección ortogonal $\bar{x}_u$ (si existe) es única. Sabemos que (si existe $\bar{x}_u$) es $\bar{x} = \bar{x}_n + (\bar{x} - \bar{x}_u)$, donde $\bar{x}_u \in U$ y $\bar{x} - \bar{x}_u \in U^\perp$; ahora bien, como la suma $U + U^\perp$ es directa (pues $U \cap U^\perp = O$; véase [144]), la anterior descomposición de $\bar{x}$ en suma de un vector de U y otro de $U^\perp$ es única, luego existe una sola proyección ortogonal $\bar{x}_u$.

EJERCICIO

En el espacio vectorial euclídeo canónico $\mathbb{R}^4$, consideremos el vector $\bar{x} = (7, 1, -6, 9)$ y el subespacio $U = \mathcal{V}(\bar{a}, \bar{b}, \bar{c})$, donde

$$\bar{a} = (1, -2, 0, 0) \quad , \quad \bar{b} = (0, 1, 2, 1) \quad , \quad \bar{c} = (-1, 0, 1, 1)$$

Se pide la proyección ortogonal de $\bar{x}$ sobre U.

RESOLUCION

Hay que descomponer $\bar{x}$ en la forma $\bar{x} = \bar{x}_u + \bar{y}$, donde $\bar{x}_u \in \bar{U}$ e $\bar{y}$ es un vector ortogonal a U. Un vector $\bar{y} = (\alpha, \beta, \gamma, \delta)$ será ortogonal a U si, y sólo si, es ortogonal a $\bar{a}$, $\bar{b}$ y $\bar{c}$, esto es, si:

$$\left.\begin{aligned} \alpha - 2\beta = 0 \\ \beta + 2\gamma + \delta = 0 \\ -\alpha + \gamma + \delta = 0 \end{aligned}\right\}$$

de donde se obtiene fácilmente que:

$$\bar{y} = (\alpha, \beta, \gamma, \delta) = \rho(2, 1, -3, 5) \quad \text{para } \rho \in \mathbb{R}$$

El subespacio U está formado por los vectores (x_1, x_2, x_3, x_4) que son ortogonales al vector hallado $(2, 1, -3, 5)$ y, por ello, tiene por ecuación a:

$$U: 2x_1 + x_2 - 3x_3 + 5x_4 = 0$$

Así, pues, hay que hallar ρ de manera que el vector

$$\bar{x}_u = \bar{x} - \rho(2, 1, -3, 5) = (7 - 2\rho, 1 - \rho, -6 + 3\rho, 9 - 5\rho)$$

pertenezca a U, o sea, de manera que

$$2(7 - 2\rho) + (1 - \rho) - 3(-6 + 3\rho) + 5(9 - 5\rho) = 0$$

de donde se obtiene que $\rho = 2$. Resulta entonces que la proyección pedida es:

$$\bar{x}_u = (3, -1, 0, -1)$$

COEFICIENTES DE FOURIER

[146]

Sea V un espacio vectorial euclídeo, en el que se consideran un subespacio $U \subset V$ y un vector $\bar{x} \in V$. Si U tiene dimensión finita y si $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es una base ortogonal de U, entonces existe la proyección ortogonal $\bar{x}_u$ de $\bar{x}$ sobre U y es:

$$\bar{x}_u = \alpha_1\bar{u}_1 + \alpha_2\bar{u}_2 + \cdots + \alpha_p\bar{u}_p \qquad [I]$$

donde

$$\alpha_i = \frac{\bar{x} \cdot \bar{u}_i}{\|\bar{u}_i\|^2} \quad (\text{para } i = 1, 2, ..., p)$$

Los escalares α_i se llaman coeficientes de Fourier de $\bar{x}$ respecto de la base ortogonal $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ de U.

DEMOSTRACION

Hemos de comprobar que, para el $\bar{x}_u$ que se obtiene de [I], se verifica que $\bar{x} - \bar{x}_u$ es ortogonal a U, es decir, que $\bar{x} - \bar{x}_u$ es ortogonal a todos los vectores de U, que es tanto como decir que $\bar{x} - \bar{x}_u$ es ortogonal a los vectores $\bar{u}_1, \bar{u}_2, ..., \bar{u}_p$, de la base dada de U. Tenemos, pues, que comprobar que $(\bar{x} - \bar{x}_u) \cdot \bar{u}_i = 0$ para $i = 1, 2, ..., p$; así ocurre ya que, como la base $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es ortogonal, se tiene:

$$(\bar{x} - \bar{x}_u) \cdot \bar{u}_i = \left(\bar{x} - \sum_{j=1}^{p} \alpha_j\bar{u}_j\right) \cdot \bar{u}_i = \bar{x} \cdot \bar{u}_i - \sum_{j=i}^{p} \alpha_j(\bar{u}_j \cdot \bar{u}_i) = \bar{x} \cdot \bar{u}_i - \alpha_i\bar{u}_i \cdot \bar{u}_i =$$

$$= \bar{x} \cdot \bar{u}_i - \frac{\bar{x} \cdot \bar{u}_i}{\|\bar{u}_i\|^2} \|\bar{u}_i\|^2 = 0$$

EJERCICIO

En un espacio vectorial euclídeo V se consideran un vector $\bar{x}$ y un sistema ortogonal de vectores no nulos $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$. Demuéstrese que

$$\|\bar{x}\|^2 \geqslant \left(\frac{\bar{x}\cdot\bar{u}_1}{\|\bar{u}_1\|}\right)^2 + \left(\frac{\bar{x}\cdot\bar{u}_2}{\|\bar{u}_2\|}\right)^2 + ... + \left(\frac{\bar{x}\cdot\bar{u}_p}{\|\bar{u}_p\|}\right)^2 \quad \text{(Desigualdad de Bessel)}$$

RESOLUCION

La proyección ortogonal de $\bar{x}$ sobre el subespacio U que engendra el sistema ortogonal $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es, según sabemos de [146]:

$$\bar{x}_u = \sum_{i=1}^{p} \alpha_i \bar{u}_i$$

donde

$$\alpha_i = \frac{\bar{x}\cdot\bar{u}_i}{\|\bar{u}_i\|^2}$$

Por tanto, como el sistema $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es ortogonal, resulta que:

$$\|\bar{x}_u\|^2 = \left\|\sum_{i=1}^{p} \alpha_i \bar{u}_i\right\|^2 = \sum_{i=1}^{p} (\alpha_i \|\bar{u}_i\|)^2 = \sum_{i=1}^{p} \left(\frac{\bar{x}\cdot\bar{u}_i}{\|\bar{u}_i\|}\right)^2$$

Por otra parte, como en la descomposición $\bar{x} = \bar{x}_u + (\bar{x} - \bar{x}_u)$ los vectores $\bar{x}_u$ y $\bar{x} - \bar{x}_u$ son ortogonales, resulta que

$$\|\bar{x}\|^2 = \|\bar{x}_u\|^2 + \|\bar{x} - \bar{x}_u\|^2 \geqslant \|\bar{x}_u\|^2 = \sum_{i=1}^{p} \left(\frac{\bar{x}\cdot\bar{u}_i}{\|\bar{u}_i\|}\right)^2$$

como había que comprobar.

γ TRANSFORMACIONES ORTOGONALES

Vamos a proceder aquí como hemos venido haciendo hasta ahora a lo largo del capítulo. Retomamos lo ya estudiado con anterioridad, en este caso lo relativo a las aplicaciones lineales, e incorporamos a ello el producto escalar. Así, nos ocuparemos de las aplicaciones lineales que «funcionan bien» en su relación con las longitudes y los ángulos, que son «compatibles» con la estructura euclídea que se añade a los espacios vectoriales. Esta compatibilidad estriba en la «conservación del producto escalar»: el producto escalar de dos vectores cualesquiera es igual al producto escalar de sus imágenes.

Tales aplicaciones, que se llaman ortogonales, no son tan novedosas como pudiera, quizá, parecer. Hay que hacer notar que en los casos de los espacios

de dimensión 2 ó 3, las susodichas transformaciones son los «movimientos» en el plano o en el espacio ordinario; estos «movimientos», convierten cualquier figura en otra que es igual a aquélla, es decir, que tiene sus mismas forma y tamaño.

8.5. TRANSFORMACIONES ORTOGONALES

APLICACIONES ORTOGONALES

[147]

Se dice que una aplicación lineal $f: V \to W$, entre los espacios vectoriales euclídeos V y W, es una *aplicación ortogonal* si conserva el producto escalar, lo que significa que:

$$f(\bar{x}) \cdot f(\bar{y}) = \bar{x} \cdot \bar{y}, \quad \forall \bar{x}, \bar{y} \in V$$

Si $f: V \to W$ es una aplicación ortogonal, entonces se verifica que

1.° f conserva las normas, esto es, $\|f(\bar{x})\| = \|\bar{x}\|$ para todo $\bar{x} \in V$.

2.° f conserva los ángulos, esto es, $(\widehat{f(\bar{x}), f(\bar{y})}) = (\widehat{\bar{x}, \bar{y}})$, $\forall \bar{x}, \bar{y} \in V$.

3.° f es una aplicación inyectiva.

La composición de dos aplicaciones ortogonales es una aplicación ortogonal.

DEMOSTRACION

1. Para cualquier vector $\bar{x} \in V$, es:

$$\|f(\bar{x})\|^2 = f(\bar{x}) \cdot f(\bar{x}) = \bar{x} \cdot \bar{x} = \|\bar{x}\|^2, \quad \text{luego } \|f(\bar{x})\| = \|\bar{x}\|$$

2. Para cualesquiera vectores $\bar{x}, \bar{y} \in V$, es:

$$\cos(f(\bar{x}), f(\bar{y})) = \frac{f(\bar{x}) \cdot f(\bar{y})}{\|f(\bar{x})\| \, \|f(\bar{y})\|} = \frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\| \, \|\bar{y}\|} = \cos(\bar{x}, \bar{y})$$

3. Dados $\bar{x}, \bar{y} \in V$, si es $f(\bar{x}) = f(\bar{y})$ entonces es $\bar{x} = \bar{y}$, ya que

$$\|\bar{x} - \bar{y}\| = \|f(\bar{x} - \bar{y})\| = \|f(\bar{x}) - f(\bar{y})\| = \|\bar{o}\| = 0$$

4. Sean $f: V \to W$ y $g: W \to U$ aplicaciones ortogonales, entre los espacios vectoriales euclídeos V, W y U. La composición $g \circ f: V \to U$ es ortogonal

ya que, además de ser lineal por serlo f y g, conserva el producto escalar, ya que para cualesquiera $\bar{x}, \bar{y} \in V$ es:

$$[(g \circ f)(\bar{x})] \cdot [(g \circ f)(\bar{y})] = g(f(\bar{x})) \cdot g(f(\bar{y})) = f(\bar{x}) \cdot f(\bar{y}) = \bar{x} \cdot \bar{y}$$

EJEMPLO

Vamos a definir una aplicación ortogonal $f: \mathbb{R}^2 \to \mathbb{R}^3$, entre los espacios vectoriales euclídeos canónicos $\mathbb{R}^2$ y $\mathbb{R}^3$. Para ello, recurramos a un sistema ortonormal $(\bar{a}, \bar{b})$ de vectores de $\mathbb{R}^3$ y tomemos

$$f(x, y) = x\bar{a} + y\bar{b} \qquad \forall (x, y) \in \mathbb{R}^2$$

Esta aplicación f, que es evidentemente lineal, conserva el producto escalar, ya que:

$$f(x, y) \cdot f(x', y') = (x\bar{a} + y\bar{b}) \cdot (x'\bar{a} + y'\bar{b}) = 1xx' + 0xy' + 0yx' + 1yy' =$$
$$= xx' + yy' = (x, y) \cdot (x', y')$$

OBSERVACIONES

1. Sea $f: V \to W$ una transformación ortogonal, entre los espacios vectoriales euclídeos V y W. De las anteriores propiedades se deduce trivialmente que: si $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es un sistema ortogonal u ortonormal de vectores de V, entonces $(f(\bar{u}_1), f(\bar{u}_2), ..., f(\bar{u}_p))$ es un sistema ortogonal u ortonormal, respectivamente, de vectores de W.
2. En la anterior definición de aplicación ortogonal se han impuesto condiciones superabundantes; pidiendo menos (aparentemente) se puede obtener el mismo resultado. En concreto, se verifica que (véanse Ejercicios III.38 y III.39):
   La aplicación $f: V \to W$, entre los espacios vectoriales euclídeos V y W, es ortogonal si se cumple una cualquiera de las tres condiciones siguientes, que son equivalentes entre sí:
   - f es lineal y conserva el producto escalar.
   - f conserva el producto escalar.
   - f es lineal y conserva la norma.
3. Según ya hemos visto, toda aplicación ortogonal es inyectiva; si, además, es sobreyectiva, recibe el nombre de *isomorfismo ortogonal*. Si $f: V \to W$ es un isomorfismo ortogonal, es evidente que la aplicación recíproca $f^{-1}: W \to V$ también es un isomorfismo ortogonal. La composición de dos isomorfismos ortogonales es, también, un isomorfismo ortogonal.
   Si $f: V \to W$ es una aplicación ortogonal y el espacio vectorial euclídeo tiene dimensión finita n, como f es inyectiva, para que sea un isomorfismo ortogonal es necesario y suficiente que W también tenga dimensión finita n.

TRANSFORMACIONES ORTOGONALES

A partir de ahora hablaremos, casi exclusivamente, de aplicaciones de un espacio V en sí mismo. Es más, nos vamos a ocupar sólo del caso de dimensión finita, que es el que aquí interesa de modo especial.

[148]

Sea $f: V \to V$ una aplicación lineal (endomorfismo), en el espacio vectorial euclídeo V. De acuerdo con la definición de aplicación ortogonal (véase [147]), se dice que f es una *transformación ortogonal* si conserva el producto escalar. Si $\dim V = n$, entonces se verifica:

1. Un endomorfismo $f: V \to V$ (con $\dim V = n$) es una transformación ortogonal si, y sólo si, la imagen por f de una base ortonormal de V es, también, base ortonormal de V.
2. Dadas dos bases ortonormales $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ y $(\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n)$ de V, existe una única transformación ortogonal $f: V \to V$ tal que $f(\bar{e}_i) = \bar{e}'_i$ (para $i = 1, 2, ..., n$).
3. Sea A la matriz de un endomorfismo $f: V \to V$ respecto de una base cualquiera $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ de V; sea G la matriz métrica del producto escalar de V en dicha base. El endomorfismo f es una transformación ortogonal si y sólo si

$$G = A^t G A$$

4. Sea A la matriz de un endomorfismo $f: V \to V$ respecto de una base ortonormal $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V. El endomorfismo f es una transformación ortogonal si y sólo si

$$A^t A = I$$

5. El conjunto de las transformaciones ortogonales en V (de dimensión finita) forman grupo, respecto de la composición de aplicaciones, que se representa por $\mathbb{O}(V)$ y recibe el nombre de *grupo ortogonal* de V.

DEMOSTRACION

1. Si f es ortogonal, como conserva ángulos y normas, transforma cualquier base ortonormal en un sistema ortonormal de n vectores, que es una base ortonormal de V, pues $\dim V = n$.
   Recíprocamente, supongamos que $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es una base ortonormal de V y que $(f(\bar{e}_1), f(\bar{e}_2), ..., f(\bar{e}_n))$ también es base ortonormal. Cualesquiera vectores $\bar{x}, \bar{y} \in V$ pueden expresarse, recurriendo a sus coordenadas, en la forma

$$\bar{x} = \sum_{i=1}^{n} x_i \bar{e}_i \quad \text{e} \quad \bar{y} = \sum_{j=1}^{n} y_j \bar{e}_j \quad (x_i, y_j \in \mathbb{R})$$

   (nótese que $\bar{x} \cdot \bar{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$)

Como f es lineal y $(f(\bar{e}_1), ..., f(\bar{e}_n))$ es base ortonormal, será:

$$f(\bar{x}) \cdot f(\bar{y}) = f(\Sigma x_i \bar{e}_i) \cdot f(\Sigma y_j \bar{e}_j) = (\Sigma x_i f(\bar{e}_i)) \cdot (\Sigma y_j f(\bar{e}_j)) =$$
$$= \Sigma\Sigma x_i y_j f(\bar{e}_i) \cdot f(\bar{e}_j) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \bar{x} \cdot \bar{y}$$

Así, pues, f conserva el producto escalar y, como es lineal, es ortogonal.

2. Según sabemos (véase [085]), hay una única aplicación lineal $f: V \to V$ que transforma la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ en la base $(\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n)$. Ahora bien, de acuerdo con la propiedad precedente, este endomorfismo f es una transformación ortogonal.
3. Denotemos por X e Y a las columnas de las coordenadas de dos vectores $\bar{x}, \bar{y} \in V$. El endomorfismo f será ortogonal si y sólo si $f(\bar{x}) \cdot f(\bar{y})$ es igual a $\bar{x} \cdot \bar{y}$ para cualesquiera $\bar{x}, \bar{y} \in V$, esto es:

$$(AX)^t G(AY) = X^t GY \quad \text{ó} \quad X^t(A^t GA)Y = X^t GY, \quad \forall \bar{x}, \bar{y} \in V$$

Esta última exigencia, por serlo para cualesquiera columnas X e Y, equivale a la igualdad $A^t GA = G$.

4. Para demostrar esta propiedad no hay más que repetir lo dicho para lo anterior, tomando ahora $G = I$ (matriz unidad), que es la matriz métrica en una base ortonormal.
5. Como todas las transformaciones ortogonales en V son automorfismos, bastará con comprobar que $\mathbb{O}(V)$ es un subgrupo del grupo lineal $GL(V)$ (véase [103]). Esto último es cierto ya que la composición de transformaciones ortogonales es transformación ortogonal, toda transformación ortogonal tiene recíproca, que también es ortogonal, y la aplicación idéntica es ortogonal.

EJERCICIO

Sea $f: \mathbb{R}^2 \to \mathbb{R}^2$ el endomorfismo definido, mediante

$$f(1, 0) = (1, -1) \quad \text{y} \quad f(0, 1) = (2, -1)$$

Sabiendo que, para cierto producto escalar de $\mathbb{R}^2$, el endomorfismo f es una transformación ortogonal, hallar la matriz métrica G de dicho producto escalar en la base canónica.

RESOLUCION

Denotemos a la matriz G como abajo se indica y tengamos en cuenta que la matriz A de el endomorfismo f en la base canónica es:

$$A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \; ; \; G = \begin{bmatrix} x & y \\ y & z \end{bmatrix}$$

La matriz G debe ser tal que $G = A^t GA$. Realizando este último producto e identificándolo con G, se obtiene:

$$\left.\begin{aligned} x &= x - 2y + z \\ y &= 2x - 3y + z \\ z &= 4x - 4y + z \end{aligned}\right\}, \text{ que equivale a } \begin{cases} x = \lambda \\ y = \lambda \\ z = 2\lambda \end{cases} \quad (\lambda \in \mathbb{R})$$

Hay, pues, infinitas soluciones, que son todas las matrices

$$G = \lambda \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \begin{pmatrix} \text{que son definidas} \\ \text{positivas para } \lambda > 0 \end{pmatrix}$$

OBSERVACIONES

1. Supongamos que $f: V \to V$ es una transformación ortogonal, definida en el espacio vectorial euclídeo V. Por ser ortogonal es lineal e inyectiva. Si V tiene dimensión finita, entonces f es biyectiva, como lo es toda aplicación lineal inyectiva entre espacios vectoriales de la misma dimensión finita (véase [084]). Ahora bien, si V no tiene dimensión finita, entonces la transformación ortogonal $f: V \to V$ pudiera no ser sobreyectiva, como prueba el siguiente ejemplo:
   Sea V el espacio vectorial de los polinomios reales. Según es fácil comprobar, la siguiente expresión de $(p(x) \mid q(x))$ define un producto escalar en el espacio V:

$$\left.\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ q(x) &= b_0 + b_1 x + \cdots + b_n x^n \end{aligned}\right\} \to (p(x) \mid q(x)) = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n$$

   (si dos polinomios son de distinto grado, añadánsele sumandos nulos al de menor grado hasta igualar al grado del otro). Sea $f: V \to V$ la aplicación definida mediante $f(p(x)) = xp(x)$; esta aplicación es evidentemente lineal y conserva el producto escalar, luego se trata de una transformación ortogonal. Pues bien, a pesar de ello, f no es sobreyectiva ya que los polinomios constantes (y no nulos) no pertenecen a la imagen de f.
2. La anterior caracterización [148], 1), de las transformaciones ortogonales, en espacio de dimensión finita, puede generalizarse fácilmente al caso en el que $f: V \to W$ es una aplicación lineal. V tiene dimensión finita y W es cualquiera. En este supuesto, se verifica que f (lineal) es una aplicación ortogonal si, y sólo si, la imagen por f de una base ortonormal de V es un sistema ortonormal de vectores de W.
   Lo mismo ocurre con la propiedad [148], 2), sobre existencia y unicidad, que se generaliza obviamente: si $(\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n)$ es una base ortonormal del espacio vectorial euclídeo V y si $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n)$ es un sistema ortonormal de vectores del espacio vectorial euclídeo W, entonces existe una única aplicación ortogonal $f: V \to W$ tal que $f(\bar{e}_i) = \bar{u}_i$ para $i = 1, 2, \ldots, n$.

3. Una transformación ortogonal $f: V \to V$ se llama *automorfismo ortogonal* si es biyectiva; si V tiene dimensión finita, toda transformación ortogonal $f: V \to V$ es automorfismo ortogonal. Las automorfismos ortogonales de un espacio vectorial euclídeo forman grupo respecto de la composición de aplicaciones.

8.6. MATRICES ORTOGONALES

Para estudiar una transformación ortogonal, las bases más idóneas son las ortonormales. En estas bases, que son las que más les cuadran a los problemas de tipo euclídeo, las ecuaciones de las transformaciones ortogonales adquieren su forma más simple. La relación $G = A^t G A$ (G = matriz métrica), que caracteriza a la matriz A de una transformación ortogonal, se reduce a $I = A^t A$ cuando la base es ortonormal, según ya se vio en [148], 3) y 4). A las matrices que verifican esta última relación se las llama ortogonales. Hablemos de ellas:

MATRICES ORTOGONALES

[149]

Sea A una matriz cuadrada de tamaño $n \times n$, de elementos reales. Se dice que A es una *matriz ortogonal* si se verifica una de las tres condiciones siguientes, que son equivalentes entre sí:

a) $A^t A = I$ (donde I = matriz unidad) o lo que es lo mismo $A^t = A^{-1}$.
b) $f_A: \mathbb{R}^n \to \mathbb{R}^n$ es una transformación ortogonal (f_A es el automorfismo que, en la base canónica, tiene a A por matriz asociada).
c) Los vectores fila (o los vectores columna) de A forman una base ortonormal del espacio vectorial euclídeo canónico $\mathbb{R}^n$.

Las matrices ortogonales de tamaño $n \times n$ forman grupo para el producto; a este grupo se le representa por $\mathbb{O}(n)$ o $\mathbb{O}_n$.

En un espacio vectorial euclídeo, al cambiar una base ortonormal por otra base ortonormal, la matriz del correspondiente cambio de coordenadas es ortogonal.

DEMOSTRACION

1. Veamos en primer lugar que las condiciones *a*), *b*) y *c*) son equivalentes. La equivalencia entre *a*) y *b*) ya está probada en el apartado anterior (véase [148], 4), por lo que solamente comprobaremos que *a*) $\Leftrightarrow$ *c*). Nótese que la relación $A^t A = I$ equivale a la $A A^t = I$, pues cualquiera de las dos signi-

fica que $A^t = A^{-1}$; por ello, en *b*) sólo se considera el caso de vectores columna (para las filas, sustitúyase A por A^t). Llamando $A = [a_{ij}]$, el vector columna de lugar i (para $i = 1, 2, ..., n$) de A es el

$$\bar{a}_i = (a_{1i}, a_{2i}, ..., a_{ni}) \in \mathbb{R}^n$$

La condición *a*), esto es, $A^tA = I$, expresada en función de los elementos a_{ij} de A, toma la forma:

$$a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (\text{para } i, j = 1, 2, ..., n)$$

Ahora bien, esta última relación significa que el producto escalar $\bar{a}_i \cdot \bar{a}_j$ (en el espacio vectorial euclídeo canónico $\mathbb{R}^n$) vale 1 ó 0 según que sea $i = j$ ó $i \neq j$, respectivamente, es decir, que $(\bar{a}_1, \bar{a}_2, ..., \bar{a}_n)$ es un sistema ortonormal de n vectores de $\mathbb{R}^n$, o sea, que se trata de una base ortonormal de $\mathbb{R}^n$, con lo que concluye la demostración.

2. Las matrices ortogonales de tamaño $n \times n$ forman grupo, para la operación producto, ya que lo forman las transformaciones ortogonales de $\mathbb{R}^n$ en $\mathbb{R}^n$ (véase [148],5), para la composición de aplicaciones. Esto es así ya que, según se sabe (véase [103]), la correspondencia $f \to A$, que a cada aplicación lineal biyectiva $f : \mathbb{R}^n \to \mathbb{R}^n$ le atribuye su matriz A en la base canónica, es un isomorfismo respecto de las operaciones antes citadas.
3. Supongamos que, desde una base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$, se pasa a la nueva base $(\bar{e}'_1, \bar{e}'_2, ..., \bar{e}'_n)$, donde

$$\bar{e}'_i = \sum_{h=1}^{n} q_{hi}\bar{e}_h \quad (i = 1, 2, ..., n)$$

Esto produce el cambio de coordenadas $X = QX'$, donde X y X' son las columnas de coordenadas en la primera y segunda bases, del mismo vector $\bar{x} \in V$, y Q es la matriz $Q = [q_{ij}]$. Si ambas bases son ortonormales, entonces

$$\bar{e}_i \cdot \bar{e}_j = \bar{e}'_i \cdot \bar{e}'_j = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases} \quad (i, j = 1, 2, ..., n)$$

y, por tanto, se verifica que:

$$\left.\begin{array}{rcl} i = j & \Rightarrow & 1 = \\ i \neq j & \Rightarrow & 0 = \end{array}\right\} \bar{e}'_i \cdot \bar{e}'_j = \left(\sum_h q_{hi}\bar{e}_h\right) \cdot \left(\sum_k q_{kj}\bar{e}_k\right) = q_{1i}q_{1j} + q_{2i}q_{2j} + \cdots + q_{ni}q_{nj}$$

es decir, las n columnas de Q forman un sistema ortonormal de vectores de $\mathbb{R}^n$, lo que significa que Q es una matriz ortogonal.

EJERCICIO

Sea A una matriz cuadrada de tamaño $n \times n$ que, descompuesta en bloques, se puede expresar en la forma:

$$A = \left[\begin{array}{c|c} B & C \\ \hline D & E \end{array}\right]$$

donde B y E son bloques cuadrados (de tamaños $p \times p$ y $q \times q$, con $p + q = n$). Sabiendo que A y B son matrices ortogonales, pruébese que ha de ser $D = O$, $C = O$ y E ortogonal.

RESOLUCION

Como $AA^t = I$ y $BB^t = I$, realizando el producto AA^t por bloques, obtenemos:

$$AA^t = \left[\begin{array}{c|c} B & C \\ \hline D & E \end{array}\right]\left[\begin{array}{c|c} B^t & D^t \\ \hline C^t & E^t \end{array}\right] = \left[\begin{array}{c|c} I + CC^t & BD^t + CE^t \\ \hline DB^t + EC^t & DD^t + EE^t \end{array}\right] = \left[\begin{array}{c|c} I & O \\ \hline O & I \end{array}\right]$$

de donde se deduce que

$$\begin{cases} I + CC^t = I & (1) \\ BD^t + CE^t = O & (2) \\ DB^t + EC^t = O & (3) \\ DD^t + EE^t = I & (4) \end{cases}$$

De (1) se deduce que $CC^t = O$ y de aquí se desprende que $C = O$, ya que, como el elemento de lugar i de CC^t es 0 (para $i = 1, 2, \ldots, n$), será:

$$0 = (c_{i1})^2 + (c_{i2})^2 + \cdots + (c_{iq})^2 \quad , \quad \text{luego } c_{i1} = c_{i2} = \cdots = c_{iq} = 0$$

Por ser $C = O$, la relación (3) queda en la forma $DB^t = O$. Como $B^tB = I$, multiplicando la anterior igualdad por B (por detrás), resulta que

$$DB^tB = OB \quad \text{luego } DI = O \quad , \quad \text{o sea } D = O$$

Como es $D = O$, la relación (4) se reduce a $EE^t = I$, luego E es ortogonal.

TRANSFORMACIONES ORTOGONALES DIRECTAS E INVERSAS

Recordemos algunas de las cosas que se estudian, en geometría elemental, al hablar de los movimientos planos; todo ello lo vamos a generalizar aquí al caso de dimensión finita n cualquiera.

Los movimientos en el plano (que son las transformaciones ortogonales en $\mathbb{R}^2$) eran de dos tipos: directos e inversos. Un movimiento era directo o inverso según que, para llevar a coincidir una figura con su transformada (manteniendo siempre su misma forma), fuera o no suficiente con desplazarla a lo largo del plano. Para lograr dicha coincidencia, en los movimientos inversos es necesario sacar la figura del plano; en los movimientos directos no se precisa de ello.

Los movimientos directos en el plano son los giros (o rotaciones); en todo movimiento inverso interviene, además de un posible giro, una simetría (respecto de una recta).

Supongamos que en el plano $\mathbb{R}^2$ se considera una base ortonormal $(\bar{e}_1, \bar{e}_2)$ con orientación positiva, es decir, tal que $\bar{e}_2$ se obtiene de girar $\bar{e}_1$ un ángulo recto en sentido positivo (contrario al de las agujas del reloj). Un movimiento transforma a $(\bar{e}_1, \bar{e}_2)$ en otra base $(\bar{e}'_1, \bar{e}'_2)$ también ortonormal. Según que el movimiento sea directo o inverso, la base $(\bar{e}'_1, \bar{e}'_2)$ tendrá orientación positiva o negativa, respectivamente. Los movimientos directos no alteran la orientación del plano; los inversos cambian dicha orientación.

ROTACIONES Y SIMETRIAS ORTOGONALES

[150]

Sea $f: V \to V$ una transformación ortogonal, en un espacio vectorial euclídeo V de dimensión finita n. Se verifica que:

1. El determinante de f, esto es(*), el determinante de cualquiera de las matrices asociada a f, vale 1 o vale -1; en uno y otro caso, se dice respectivamente que f es una *transformación ortogonal directa o inversa*. Las primeras, que también se llaman *rotaciones*, forman grupo (grupo de las rotaciones), que se denota por $\mathbb{O}^+(V)$.
2. Si $\det f = -1$ (f es ortogonal inversa), entonces f se compone de una *simetría ortogonal*, respecto de un subespacio de dimensión $n-1$(**), y de una rotación (que puede ser la identidad); es decir, $f = (\text{simetría}) \circ (\text{rotación})$.
3. Una base de V y su base transformada por f tienen la misma o distinta orientación(***) según que f sea, respectivamente, directa o inversa.

(*) Según se vio en [095], todas las matrices asociadas a un mismo endomorfismo tienen el mismo determinante.

(**) Si U es un subespacio de V, llamando $\bar{x}_u$ a la proyección ortogonal de un vector $\bar{x} \in V$ sobre U, la simetría ortogonal respecto de U es la aplicación $\bar{x} \to \bar{x}'$ definida por $\bar{x} + \bar{x}' = 2\bar{x}_u$.

(***) Véase el concepto de bases igualmente orientadas en [095].

DEMOSTRACION

1.° Antes de empezar, quizá convenga recordar que si A y A' son matrices de f en bases distintas, como $A' = P^{-1}AP$, donde P es la matriz regular del cambio de coordenadas, tomando determinantes en esta igualdad, como $\det P^{-1} = (\det P)^{-1}$, resulta que $\det A' = \det A$. Una vez hecha esta comprobación previa, que nos permite poner $\det f = \det A = \det A'$, veamos ya que $\det f = 1$ o $\det f = -1$, o lo que es igual, que $(\det f)^2 = 1$.

Recurramos a que, si A es la matriz de f en una cierta base de V y llamando G a la matriz métrica del producto escalar en esa base, se verifica que $G = A^t GA$. Tomando aquí determinante, como G es una matriz regular y como $\det A^t = \det A$, se obtiene:

$$\det G = (\det A^t)(\det G)(\det A), \quad \text{luego} \quad 1 = (\det A)^2$$

Es decir $(\det f)^2 = 1$, como había que comprobar.

El conjunto $\mathbb{O}^+(V)$, de las rotaciones en V. es un subgrupo del grupo ortogonal $\mathbb{O}(V)$ y ello es debido a que el determinante de la composición de dos endomorfismos es igual al producto de los determinantes de los endomorfismos componentes. En efecto: 1) La identidad es una rotación, pues $\det I = 1$; 2) la composición de dos rotaciones f_1 y f_2 es otra rotación, ya que $\det(f_1 \circ f_2) = (\det f_1)(\det f_2) = 1 \times 1 = 1$; 3) Si f es una rotación, entonces f^{-1} también es rotación, ya que $(\det f^{-1}) = (\det f)^{-1} = = (1)^{-1} = 1$.

2.° Hemos de comprobar que si $\det f = -1$, entonces f es la composición de una simetría ortogonal (respecto de un subespacio de dimensión $n-1$) por una rotación (o transformación ortogonal directa). Para ello, echemos mano de una base ortonormal $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V; llamemos A a la matriz de f en esta base, que es una matriz ortogonal inversa, es decir, tal que $A^tA = I$ y $\det A = -1$. Sea A_1 la matriz que se obtiene de cambiar de signo a todos los elementos de la primera fila de A, sin ninguna otra alteración. Sea I_1 la matriz que sólo difiere de la matriz unidad I en que su elemento de lugar 11 es -1 (en lugar de 1, como es en I). Es evidente que $A = I_1A_1$; como consecuencia de ello, se verifica que $\det A_1 = 1$, ya que $\det A = -1$ y $\det I_1 = -1$. Por otra parte, como $I_1^tI_1 = I$, resulta que

$$A^tA = I \;\Rightarrow\; (I_1A_1)^t(I_1A_1) = I \;\Rightarrow\; A_1^t(I_1^tI_1)A_1 = I \;\Rightarrow\; A_1^tA_1 = I$$

luego A_1 es ortogonal y, además, es ortogonal directa, pues $\det A_1 = 1$.

Como A_1 es ortogonal directa, también será ortogonal directa la transformación $f_1 : V \to V$ que tiene asociada la matriz A_1, en la base dada; es decir, f_1 es una rotación. La transformación $s : V \to V$ que, en la referida base, tiene por matriz a la I_1 es la simetría ortogonal respecto del subespacio engendrado por $(\bar{e}_2, ..., \bar{e}_n)$. Al ser $A = I_1A_1$, se concluye que $f = s \circ f_1$, como se deseaba obtener.

3.° En este apartado, lo único que se hace es particularizar lo ya dicho anteriormente (véase [095]) sobre la conservación o no conservación de la orientación de las bases por un automorfismo; aquí, los automorfismos que se consideran son las transformaciones ortogonales.

EJERCICIO

En el espacio vectorial euclídeo canónico $\mathbb{R}^3$, sea $f:\mathbb{R}^3\to\mathbb{R}^3$ un giro (o rotación), de cierto ángulo, alrededor del vector $\bar{u}=(1,1,1)$. Compruébese que la matriz A asociada a f en la base canónica es del tipo siguiente:

$$A=\begin{bmatrix} a & c & b \\ b & a & c \\ c & b & a \end{bmatrix} \quad \text{con} \begin{cases} a^2+b^2+c^2=1 \\ a+b+c=1 \end{cases}$$

RESOLUCION

Como el vector $\bar{u}$ forma el mismo ángulo con los tres vectores de la base, si llamamos (a, b, c) a f$(1, 0, 0)$, por simetría de revolución las imágenes de los otros dos vectores de la base serán $f(0, 1, 0)=(c, a, b)$ y $f(0, 0, 1)=(b, c, a)$. Las columnas de A son, por tanto, las que se señalan en el enunciado. Como dichas columnas son vectores unitarios de $\mathbb{R}^n$, ha de ser $a^2+b^2+c^2=1$. Como $\bar{u}$ permanece invariable y f conserva el producto escalar, si es $\bar{e}_1=(1, 0, 0)$, ha de ser:

$$f(\bar{e}_1)\cdot f(\bar{u})=\bar{e}_1\cdot\bar{u} \quad , \quad \text{o sea } (a, b, c)\cdot(1, 1, 1)=(1, 0, 0)\cdot(1, 1, 1)$$

es decir, debe de ser $a+b+c=1$.

Nótese que las columnas de A son ortogonales dos a dos, ya que

$$1^2=(a+b+c)^2=(a^2+b^2+c^2)+2(ab+bc+ca)=1+2(ab+bc+ca)$$

luego $ab+bc+ca=0$, lo que confirma la ortogonalidad de las columnas de A.

MATRICES ORTOGONALES DIRECTAS E INVERSAS

[151]

Sea A una matriz ortogonal, de tamaño $n\times n$. Se verifica que:

1. El determinante de A vale $\det A=1$ o vale $\det A=-1$; se dice, respectivamente, que A es *matriz ortogonal directa* o *inversa*.
2. Las matrices ortogonales directas de tamaño $n\times n$ forman grupo, al que se denota por $\mathbb{O}^+(n)$ o $\mathbb{O}_n^+$.

DEMOSTRACION

1. Aunque podemos remitirnos a que el determinante de una transformación ortogonal vale ± 1, demostremos directamente esta propiedad: como A es ortogonal, será $A^t A = I$, y tomando determinantes se obtiene:

$$(\det A^t)(\det A) = 1 \quad , \quad \text{luego } (\det A)^2 = 1$$

es decir

$$\det A = \pm 1$$

2. El conjunto $\mathbb{O}^+(n)$ es un subgrupo del grupo $\mathbb{O}(n)$ de las matrices ortogonalesde tamaño $n \times n$. En efecto:

   - Si I es la matriz unidad, $I \in \mathbb{O}^+(n)$, ya que $I \in \mathbb{O}(n)$ y $\det I = 1$.
   - $A, B \in \mathbb{O}^+(n) \Rightarrow AB^{-1} \in \mathbb{O}^+(n)$, ya que $AB^{-1} \in \mathbb{O}(n)$ y $\det(AB^{-1}) = (\det A)(\det B)^{-1} = 1 \times 1^{-1} = 1$.

EJERCICIO

Compruébese que la matriz A del ejercicio anterior (que, según allí se vio, es ortogonal) es ortogonal directa.

RESOLUCION

Sumándole a la última columna las dos primeras, como $a + b + c = 1$, se obtiene:

$$\det A = \begin{vmatrix} a & c & 1 \\ b & a & 1 \\ c & b & 1 \end{vmatrix} = (a^2 + b^2 + c^2) - (ac + ba + cb) = 1 - 0 = 1$$

8.7. TRANSFORMACIONES ORTOGONALES EN 2 Y 3 DIMENSIONES

No se pretende hacer aquí un estudio pormenorizado del funcionamiento de las transformaciones ortogonales en espacios euclídeos de dimensión n cualquiera, para el que, entre otras cosas, precisaríamos de algunos conocimientos acerca de los autovalores y de los autovectores de las matrices, que aún no poseemos. Nos vamos a ceñir a los casos de dimensiones $n = 2$ y $n = 3$. Más adelante (véase [183] a [185]) analizaremos el caso general (n cualquiera).

El análisis que vamos a hacer aquí de estas cuestiones no es exhaustivo, se limita a recoger, comprobándolos, sus aspectos más destacados.

TRANSFORMACIONES ORTOGONALES EN DIMENSION 2

[152]

1.° Sea $f: V_2 \to V_2$ un movimiento (transformación ortogonal), en un espacio vectorial euclídeo V_2, de dimensión 2. Según que f sea un movimiento directo (rotación) o inverso (simetría), la matriz de f en una base ortonormal $(\bar{e}_1, \bar{e}_2)$ de V_2 es la siguiente A_R o A_S, respectivamente, para algún $\alpha \in]-\pi, \pi]$.

2.° MATRIZ DE ROTACION: La matriz (en base ortonormal) del movimiento directo (o rotación) f es la siguiente A_R (donde α no depende de la base; α sólo depende de f):

$$A_R = \begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

(en un caso general) (si $\alpha = 0$) (si $\alpha = \pi$)

El movimiento directo cuya matriz es A_R es la rotación de ángulo α en sentido positivo (esto es, de $\bar{e}_1$ a $\bar{e}_2$) si $0 < \alpha < \pi$ y en sentido negativo si $-\pi < \alpha < 0$. En cualquier otra base ortonormal con la misma orientación que la de $(\bar{e}_1, \bar{e}_2)$, la matriz de f no cambia, sigue siendo la misma A_R.

3.° MATRIZ DE SIMETRIA ORTOGONAL: La matriz (en base ortonormal) del movimiento inverso (o simetría ortogonal) f es la siguiente A_S (donde α depende de la base ortonormal, y sólo de ella):

$$A_S = \begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{bmatrix} \qquad A_{S_0} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

(en base ortonormal cualquiera) (en base ortonormal ad hoc)

El movimiento inverso cuya matriz es A_S es la simetría ortogonal respecto de la recta E (eje) engendrada por el vector $\bar{a}$ de coordenadas $(\cos(\alpha/2), \operatorname{sen}(\alpha/2))$. La matriz A_{S_0} es la matriz de f en una base ortonormal cuyo primer vector es $\bar{a}$.

COMPROBACION

1. Sea A la matriz de f en la base dada y representémosla poniendo:

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Como A es ortogonal, ha de ser $a^2 + b^2 = 1$, luego existe un único $\alpha \in]-\pi, \pi]$ tal que $a = \cos\alpha$ y $b = \operatorname{sen}\alpha$. Como A es ortogonal, ha de ser $a^2 + c^2 = 1$, luego $c^2 = \operatorname{sen}^2\alpha$ y por ello $c = \pm\operatorname{sen}\alpha$ ó $c = \varepsilon\operatorname{sen}\alpha$,

donde $\varepsilon = \pm 1$. Como A es ortogonal, ha de ser $ab + cd = 0$, luego $d = -\varepsilon \cos \alpha$(*). Por tanto $\det A = -\varepsilon$ y, como $\det A$ vale 1 si f es directa y -1 si f es inversa, resulta que $\varepsilon = -1$ ó $\varepsilon = 1$ según que f sea directa o inversa, respectivamente, con lo que concluye la comprobación.

2. Si la matriz de f es A_R, entonces los transformados de los vectores $\bar{e}_1$ y $\bar{e}_2$ de la base son los $\bar{e}'_1$ y $\bar{e}'_2$ siguientes:

$$\bar{e}_1 \mapsto \bar{e}'_1 = \cos \alpha\, \bar{e}_1 + \operatorname{sen} \alpha\, \bar{e}_2 \quad \text{y} \quad \bar{e}_2 \mapsto \bar{e}'_2 = -\operatorname{sen} \alpha\, \bar{e}_1 + \cos \alpha\, \bar{e}_2$$

Nótese que $\bar{e}'_1$ y $\bar{e}'_2$ son los vectores que resultan de rotar $\bar{e}_1$ y $\bar{e}_2$ un ángulo α (en el sentido que se señala en el enunciado). Como (por ser f lineal) el transformado de un vector $\bar{u} = x\bar{e}_1 + y\bar{e}_2$, cualquiera de V_2, es

$$\bar{u} = x\bar{e}_1 + y\bar{e}_2 \mapsto \bar{u}' = x\bar{e}'_1 + y\bar{e}'_2$$

resulta que cualquier vector de V_2 se transforma rotándole un ángulo α, del modo que se dice en el enunciado.

Sea A' la matriz de f en otra base ortonormal $(\bar{u}_1, \bar{u}_2)$ con la misma orientación que $(\bar{e}_1, \bar{e}_2)$. Como f es el giro de ángulo α, resulta que

$$f(\bar{u}_1) = \cos \alpha\, \bar{u}_1 + \operatorname{sen} \alpha\, \bar{u}_2 \quad \text{y} \quad f(\bar{u}_2) = -\operatorname{sen} \alpha\, \bar{u}_1 + \cos \alpha\, \bar{u}_2$$

Por tanto, las columnas de A' son (1.ª columna) = $(\cos \alpha, \operatorname{sen} \alpha)$ y (2.ª columna) = $(-\operatorname{sen} \alpha, \cos \alpha)$, es decir, A' coincide con la anterior matriz A_R.

3. Si la matriz de f es A_S, llamemos $\bar{a}$ y $\bar{b}$ a los siguientes vectores unitarios

$$\bar{a} = \cos \frac{\alpha}{2}\, \bar{e}_1 + \operatorname{sen} \frac{\alpha}{2} \bar{e}_2 \qquad \text{y} \qquad \bar{b} = \operatorname{sen} \frac{\alpha}{2}\, \bar{e}_1 - \cos \frac{\alpha}{2}\, \bar{e}_2$$

Recuérdese que se llamó E a la recta vectorial que determina el vector $\bar{a}$. Los transformados de estos vectores $\bar{a}$ y $\bar{b}$ son los siguientes $\bar{a}'$ y $\bar{b}'$:

$$\bar{a}' = \left(\cos \alpha \cos \frac{\alpha}{2} + \operatorname{sen} \alpha \operatorname{sen} \frac{\alpha}{2}\right)\bar{e}_1 + \left(\operatorname{sen} \alpha \cos \frac{\alpha}{2} - \cos \alpha \operatorname{sen} \frac{\alpha}{2}\right)\bar{e}_2 =$$

$$= \cos \frac{\alpha}{2}\, \bar{e}_1 + \operatorname{sen} \frac{\alpha}{2}\, \bar{e}_2 = \bar{a}$$

$$\bar{b}' = \left(\cos \alpha \operatorname{sen} \frac{\alpha}{2} - \operatorname{sen} \alpha \cos \frac{\alpha}{2}\right)\bar{e}_1 + \left(\operatorname{sen} \alpha \operatorname{sen} \frac{\alpha}{2} + \cos \alpha \cos \frac{\alpha}{2}\right)\bar{e}_2 =$$

$$= -\operatorname{sen} \frac{\alpha}{2}\, \bar{e}_1 + \cos \frac{\alpha}{2}\, \bar{e}_2 = -\bar{b}$$

Por tanto, como cualquier vector $\bar{u} \in V_2$ se puede expresar en la forma $\bar{u} = x\bar{a} + y\bar{b}$ (nótese que $\bar{a}$ y $\bar{b}$ forman base ortonormal de V_2) y recurriendo

(*) Si fuese $c = 0$, como $c^2 + d^2 = 1$, resultaría que $d = \pm 1$; como ahora sería $b = 0$ y $a = \pm 1$, seguía siendo válida la conclusión $d = -\varepsilon \cos \alpha$.

a que f es lineal, resulta que el transformado de cualquier $\bar{u} \in V_2$ es:

$$x\bar{a} + y\bar{b} \longmapsto f(x\bar{a} + y\bar{b}) = x\bar{a}' + y\bar{b}' = x\bar{a} - y\bar{b}$$

Como $x\bar{a} - y\bar{b}$ es el simétrico de $x\bar{a} + y\bar{b}$ respecto de la recta E, la propiedad queda comprobada.

Como $f(\bar{a}) = \bar{a}$ y $f(\bar{b}) = -b$, la matriz de f en la base ortonormal $(\bar{a}, \bar{b})$ tiene (1.ª columna) = (1, 0) y (2.ª columna) = (0, −1), es decir, esta matriz es la A_{S_0} del enunciado.

EJERCICIO

Sea $f: V_2 \to V_2$ la transformación ortogonal, en el espacio vectorial euclídeo bidimensional V_2, que respecto de una base ortonormal $(\bar{e}_1, \bar{e}_2)$ de V_2 tiene asociada la siguiente matriz

$$A = \begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{bmatrix}$$

Obtener directamente aquellos vectores $\bar{u}, \bar{v} \in V_2$ tales que $f(\bar{u}) = \bar{u}$ y $f(\bar{v}) = -\bar{v}$.

RESOLUCION

Llamando x e y a las coordenadas de los vectores buscados, tanto de $\bar{u}$ como de $\bar{v}$, se ha de verificar que:

$$\begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \varepsilon \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{donde} \quad \varepsilon = \begin{cases} 1 \text{ para el vector } \bar{u} \\ -1 \text{ para el vector } \bar{v} \end{cases}$$

Operando en la anterior ecuación matricial, se obtiene:

$$\left.\begin{aligned} (\cos\alpha - \varepsilon)x + \operatorname{sen}\alpha\, y = 0 \\ \operatorname{sen}\alpha\, x - (\cos\alpha + \varepsilon)y = 0 \end{aligned}\right\}$$

donde

$$\begin{vmatrix} \cos\alpha - \varepsilon & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha + \varepsilon \end{vmatrix} = \varepsilon^2 - 1 = 0$$

Este sistema tiene pues solución, que es $x = \operatorname{sen}\alpha$, $y = \varepsilon - \cos\alpha$ o cualquier otra proporcional a ella. Recurriendo al ángulo mitad $\alpha/2$ y poniendo $\operatorname{sen}\alpha = 2\operatorname{sen}(\alpha/2)\cos(\alpha/2)$ y $\cos\alpha = 2\cos^2(\alpha/2) - 1 = 1 - 2\operatorname{sen}^2(\alpha/2)$, se obtiene fácilmente que

$$\bar{u} = \rho\left(\cos\frac{\alpha}{2}\,\bar{e}_1 + \operatorname{sen}\frac{\alpha}{2}\,\bar{e}_2\right) \quad \text{y} \quad \bar{v} = \rho\left(\operatorname{sen}\frac{\alpha}{2}\,\bar{e}_1 - \cos\frac{\alpha}{2}\,\bar{e}_2\right), \quad \text{para } \rho \in \mathbb{R}$$

TRANSFORMACIONES ORTOGONALES EN DIMENSION 3

[153]

Sea $f: V_3 \to V_3$ un movimiento (transformación ortogonal), en un espacio vectorial euclídeo V_3 de dimensión 3. Se verifica que:

1. Existe una base ortonormal $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ de V_3 en la que la matriz (ortonormal) de f es una de las $A_\varepsilon(\alpha)$ siguientes (donde $\varepsilon = 1$ ó $\varepsilon = -1$), para cierto $\alpha \in]-\pi, \pi]$ (*):

$$A_\varepsilon(\alpha) = \begin{bmatrix} \varepsilon & 0 & 0 \\ 0 & \cos\alpha & -\operatorname{sen}\alpha \\ 0 & \operatorname{sen}\alpha & \cos\alpha \end{bmatrix} \quad \begin{pmatrix} \varepsilon = 1 & \text{si } f \text{ es directo} \\ \varepsilon = -1 & \text{si } f \text{ es inverso} \end{pmatrix}$$

Casos particulares de especial relevancia ($\alpha = 0$ y $\alpha = \pi$)

$$A_1(0) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad ; \quad A_1(\pi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(identidad) (simetría respecto de una recta)

$$A_{-1}(0) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad ; \quad A_{-1}(\pi) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(simetría respecto de un plano) (simetría respecto del origen)

2. Si la matriz de f es $A_{+1}(\alpha)$, entonces f es la rotación de ángulo α alrededor de la recta R engendrada por el vector $\bar{e}_1$. En particular, si $\alpha = 0$ entonces f es la identidad y si $\alpha = \pi$ entonces f es la simetría respecto de R.
3. Si la matriz de f es $A_{-1}(\alpha)$, entonces f es la composición de la rotación del apartado anterior con la simetría ortogonal respecto del plano P ortogonal a R, esto es, el que engendran los vectores $\bar{e}_2$ y $\bar{e}_3$. En particular, si $\alpha = 0$, entonces f es la simetría respecto de P y si $\alpha = \pi$ entonces f es la simetría respecto del origen (vector nulo).

(*) Para probar esta afirmación se va a recurrir a un resultado que se comprobará posteriormente (cuando se estudien los autovalores de endomorfismos y de matrices). Se trata de lo siguiente: si f es un endomorfismo ortogonal en un espacio tridimensional, entonces existe algún vector $\bar{u} \neq \bar{0}$ tal que $f(\bar{u}) = \bar{u}$ o tal que $f(\bar{u}) = -\bar{u}$ (véase [180]).

COMPROBACION

1. Según hemos anunciado, se va a echar mano de un resultado que más adelante se podrá comprobar con facilidad, con el recurso de los autovalores y los vectores propios de los endomorfismos. Debido a que el endomorfismo f es ortogonal y como la dimensión de V_3 es impar, posteriormente se obtendrá fácilmente que ha de existir algún vector $u \in V_3$ (no nulo) tal que $f(\bar{u}) = \bar{u}$ ó $f(\bar{u}) = -\bar{u}$ (véase [180],1). Sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ una base ortonormal de V_3 a la que sólo se la exige que $\bar{e}_1$ sea vector de la dirección del anterior vector $\bar{u}$, con lo que $f(\bar{e}_1) = \pm\bar{e}_1$; sea A la matriz de f en esta base. La primera columna de A, como está formada por las coordenadas de $f(\bar{e}_1) = \pm\bar{e}_1$, es la $(a_{11}, a_{21}, a_{31}) = (\pm 1, 0, 0)$. Como A es ortogonal, su primera fila $(\pm 1, a_{12}, a_{13})$ es un vector unitario de $\mathbb{R}^3$ y, por ello, ha de ser $a_{12} = a_{13} = 0$. De lo ya dicho y como A es ortogonal, resulta que:

$$A = \begin{bmatrix} \pm 1 & 0 & 0 \\ 0 & a & c \\ 0 & b & d \end{bmatrix}$$

verificándose:

$$\begin{cases} 0^2 + a^2 + b^2 = 1 \\ 0^2 + c^2 + d^2 = 1 \\ 0^2 + ac + bd = 0 \end{cases}$$

es decir, descomponiendo A en bloques

$$A = \left[\begin{array}{c|cc} \varepsilon & 0 & 0 \\ \hline 0 & & \\ 0 & & B \end{array}\right]$$

donde

$$\begin{cases} \varepsilon = 1 \text{ ó } \varepsilon = -1 \\ B \text{ es una matriz ortogonal cualquiera de tamaño } 2 \times 2 \end{cases}$$

Si B es ortogonal directa, entonces sus columnas son, según se ha comprobado en [152] (1.ª columna) = $(\cos\alpha, \operatorname{sen}\alpha)$ y (2.ª columna) = $(-\operatorname{sen}\alpha, \cos\alpha)$ para un cierto $\alpha \in]-\pi, \pi]$. Por tanto, si B es ortogonal directa, en cualquiera de las bases $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ que venimos manejando la matriz de f es una de las $A_\varepsilon(\alpha)$ del enunciado.

Si B es ortogonal inversa, entonces sus columnas son, según obtuvimos en [152], (1.ª columna) = $(\cos\alpha, \operatorname{sen}\alpha)$ y (2.ªcolumna) = $(\operatorname{sen}\alpha, -\cos\alpha)$ para algún $\alpha \in]-\pi, \pi]$, que dependerá de la base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ que venimos considerando. En este caso, al cambiar a la nueva base ortonormal $(\bar{e}_1, \bar{a}, \bar{b})$, donde

$$\bar{a} = \cos\frac{\alpha}{2}\bar{e}_2 + \operatorname{sen}\frac{\alpha}{2}\bar{e}_3 \quad \text{y} \quad \bar{b} = \operatorname{sen}\frac{\alpha}{2}\bar{e}_2 - \cos\frac{\alpha}{2}\bar{e}_3$$

la matriz de f pasa a ser la

$$\begin{bmatrix} \varepsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(esto se comprueba, como en el caso de dimensión 2, viendo que $f(\bar{a}) = \bar{a}$ y que $f(\bar{b}) = -b$). Por tanto: para $\varepsilon = +1$, la matriz de f en la base $(\bar{b}, \bar{a}, \bar{e}_1)$ es la $A_{-1}(0)$ del enunciado; y para $\varepsilon = -1$, la matriz de f en la base $(\bar{a}, \bar{b}, \bar{e}_1)$ es la $A_1(\pi)$ del enunciado.

2. Para $\varepsilon = +1$, los transformados de los vectores $\bar{e}_1$, $\bar{e}_2$ y $\bar{e}_3$ de la base son

$$\bar{e}_1' = \bar{e}_1 \quad , \quad \bar{e}_2' = \cos\alpha\,\bar{e}_2 + \operatorname{sen}\alpha\,\bar{e}_3 \quad \text{y} \quad \bar{e}_3' = -\operatorname{sen}\alpha\,\bar{e}_2 + \cos\alpha\,\bar{e}_3$$

Estos vectores son, como se ve fácilmente, los que resultan de rotar los $\bar{e}_1$, $\bar{e}_2$ y $\bar{e}_3$ un ángulo α alrededor de la recta E (engendrada por $\bar{e}_1$). El transformado de un vector cualquiera $\bar{u} \in V_3$ será (por ser f lineal)

$$\bar{u} = x_1\bar{e}_1 + x_2\bar{e}_2 + x_3\bar{e}_3 \longmapsto \bar{u}' = x_1\bar{e}_1' + x_2\bar{e}_2' + x_3\bar{e}_3'$$

que, por tanto, es el vector que resulta de rotar $\bar{u}$ el ángulo α alrededor de la recta E.

3. Para $\varepsilon = -1$, la matriz $A_\varepsilon(\alpha)$ se puede poner:

$$A_{-1}(\alpha) = A_1(\alpha) \cdot A_{-1}(0)$$

Como $A_{-1}(\pi)$ es la matriz de la simetría ortogonal respecto del plano P (engendrado por $\bar{e}_2$ y $\bar{e}_3$) y $A_1(\alpha)$ es la matriz de la rotación de ángulo α alrededor de R (recta engendrada por $\bar{e}_1$), resulta que $A_{-1}(\alpha)$ es la matriz de la composición de las dos transformaciones citadas, como debíamos comprobar.

EJERCICIO

Sea $f: \mathbb{R}^3 \to \mathbb{R}^3$ el endomorfismo que en la base canónica tiene asociada la matriz

$$A = \begin{bmatrix} \dfrac{\sqrt{2}-2}{4} & \dfrac{-\sqrt{2}-2}{4} & \dfrac{1}{2} \\ \dfrac{-\sqrt{2}-2}{4} & \dfrac{\sqrt{2}-2}{4} & -\dfrac{1}{2} \\ -\dfrac{1}{2} & \dfrac{1}{2} & \dfrac{\sqrt{2}}{2} \end{bmatrix}$$

La matriz A es ortogonal (como se comprueba fácilmente). Averígüese el tipo de movimiento que es f y describirlo en términos de rotaciones y simetrías.

RESOLUCION

Resolvamos la ecuación $f(\bar{x}) = \varepsilon\bar{x}$ (para $\varepsilon = \pm 1$; $\bar{x} \in \mathbb{R}^3$). Llamando X a columna de coordenadas del vector incógnita $\bar{x}$, la ecuación a resolver es $AX = \varepsilon X$, que se puede poner $AX = \varepsilon I\bar{X}$ (I matriz unidad) o también $(A - \varepsilon I)X = O$, que es un sistema homogéneo cuya matriz es $A - \varepsilon I$. Como se buscan soluciones no nulas de este sistema, ha de ser $\det(A - \varepsilon I) = 0$. Es evidente que esto es así para $\varepsilon = -1$, pues las dos primeras filas de $A + I$ son proporcionales. Para $\varepsilon = -1$, la ecuación $(A - \varepsilon I)X = O$ conduce a:

$$\left.\begin{aligned}(\sqrt{2}+2)x_1 - (\sqrt{2}+2)x_2 + 2x_3 = 0\\ -x_1 + x_2 + \sqrt{2}x_3 = 0\end{aligned}\right\} \quad , \quad \text{que dé} \begin{cases} x_1 = \rho \\ x_2 = \rho \ (\rho \in \mathbb{R}) \\ x_3 = 0 \end{cases}$$

Por tanto, el vector $\bar{x} = (1, 1, 0)$ es tal que $f(\bar{x}) = -\bar{x}$. Tomemos una nueva base ortonormal $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ en la que $\bar{e}_1$, sea unitario de la dirección de $\bar{x}$; por ejemplo:

$$\bar{e}_1 = (1/\sqrt{2}, 1/\sqrt{2}, 0) \quad , \quad \bar{e}_2(-1/\sqrt{2}, 1/\sqrt{2}, 0) \quad \text{y} \quad \bar{e}_3 = (0, 0, 1)$$

La matriz A' de f en la nueva base es $A' = P^{-1}AP$, donde P es la matriz de cambio de coordenadas, es decir:

$$P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad P^{-1} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{y} \quad A' = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Por tanto, f es la composición de la rotación de 45° alrededor de la recta engendrada por $\bar{e}_1 = (1/\sqrt{2}, 1/\sqrt{2}, 0)$ y la simetría ortogonal respecto del plano ortogonal a dicha recta.

δ PRODUCTO MIXTO Y PRODUCTO VECTORIAL

Aun cuando lo que aquí vamos a decir acerca del producto mixto y del producto vectorial es fácilmente generalizable al caso de espacios vectoriales euclídeos de dimensión finita n cualquiera, nos vamos a centrar en el caso $n = 3$, que es para el que vamos a necesitar, más adelante, de estos conceptos. Ello no obstante, se harán algunas observaciones y comentarios sobre el caso general (n cualquiera).

8.8. PRODUCTO MIXTO

DEFINICION

[154]

Sea V_3 un espacio vectorial euclídeo, orientado(*) de dimensión 3, y sea $(\bar{x}, \bar{y}, \bar{z})$ una terna de vectores de V_3. Si las coordenadas de estos vectores en una base ortonormal directa son $\bar{x}(x_1, x_2, x_3)$, $\bar{y}(y_1, y_2, y_3)$ y $\bar{z}(z_1, z_2, z_3)$, entonces el siguiente determinante, que denotaremos poniendo $[\bar{x}, \bar{y}, \bar{z}]$:

$$[\bar{x}, \bar{y}, \bar{z}] = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

no depende de la base (ortonormal directa) que se tome en V_3. Este número $[\bar{x}, \bar{y}, \bar{z}]$, que sólo depende de $(\bar{x}, \bar{y}, \bar{z})$, se llama *producto mixto* de esta terna de vectores.

(*) Recuérdese (véase [095]) que, atendiendo a su orientación, las bases de un espacio vectorial se dividían en dos clases. Se conviene en considerar que las bases que forman una de estas dos clases tienen orientación positiva o, dicho de otro modo, que son bases directas. En el espacio vectorial $\mathbb{R}^3$, se suele considerar que la base canónica es directa.

DEMOSTRACION

Llamemos X, Y y Z, a las columnas de coordenadas de $\bar{x}$, $\bar{y}$, y $\bar{z}$ en una base ortonormal directa; llamemos X', Y' y Z', a las columnas de coordenadas de $\bar{x}$, $\bar{y}$ y $\bar{z}$ en otra base ortonormal directa; sea P la matriz de cambio de coordenadas, de manera que

$$X = PX' \quad , \quad Y = PY' \quad \text{y} \quad Z = PZ' \qquad [1]$$

La matriz P es ortogonal directa, ya que las dos bases son ortonormales y tienen la misma orientación; por tanto, $\det P = 1$. Como las relaciones [1] se pueden expresar conjuntamente recurriendo a la matriz $[XYZ]$ cuyas columnas son X, Y y Z, poniendo

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} = [XYZ] = P[X'Y'Z'] = P \begin{bmatrix} x'_1 & y'_1 & z'_1 \\ x'_2 & y'_2 & z'_2 \\ x'_3 & y'_3 & z'_3 \end{bmatrix}$$

tomando determinantes en ambos miembros de esta igualdad, se obtiene:

$$[\vec{x}, \vec{y}, \vec{z}] = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = (\det P) \begin{vmatrix} x'_1 & y'_1 & z'_1 \\ x'_2 & y'_2 & z'_2 \\ x'_3 & y'_3 & z'_3 \end{vmatrix} = 1 \cdot [\vec{x}, \vec{y}, \vec{z}]$$

en la 1.ª base — en la 2.ª base

PROPIEDADES

[155]

Sea V_3 un espacio vectorial euclídeo, orientado, de dimensión 3. El *producto mixto*, que también es la aplicación que a cada terna de vectores $(\vec{x}, \vec{y}, \vec{z})$ de V_3 le atribuye su producto mixto $[\vec{x}, \vec{y}, \vec{z}]$, es decir:

$$V_3 \times V_3 \times V_3 \to \mathbb{R} \quad , \quad (\vec{x}, \vec{y}, \vec{z}) \mapsto [\vec{x}, \vec{y}, \vec{z}]$$

goza de las siguientes propiedades. Para cualquiera que sean los vectores $\vec{x}, \vec{x}', \vec{y}, \vec{y}', \vec{z}$ y $\vec{z}'$ de V_3 y los escalares $\lambda, \lambda', \mu, \mu', \nu, \nu' \in \mathbb{R}$ se verifica que:

- El producto mixto es una forma lineal respecto de cada una de sus tres variables (trilineal), es decir:

$$[\lambda\vec{x} + \lambda'\vec{x}', \mu\vec{y} + \mu'\vec{y}', \nu\vec{z} + \nu'\vec{z}'] = \lambda\mu\nu[\vec{x}, \vec{y}, \vec{z}] + \lambda\mu\nu'[\vec{x}, \vec{y}, \vec{z}'] + \\ + \lambda\mu'\nu[\vec{x}, \vec{y}', \vec{z}] + \lambda\mu'\nu'[\vec{x}, \vec{y}', \vec{z}'] + \lambda'\mu\nu[\vec{x}', \vec{y}, \vec{z}] + \\ + \lambda'\mu\nu'[\vec{x}', \vec{y}, \vec{z}'] + \lambda'\mu'\nu[\vec{x}', \vec{y}', \vec{z}] + \lambda'\mu'\nu'[\vec{x}', \vec{y}', \vec{z}']$$

- El producto mixto es una forma antisimétrica, es decir:

$$[\vec{x}, \vec{y}, \vec{z}] = -[\vec{x}, \vec{z}, \vec{y}] = [\vec{z}, \vec{x}, \vec{y}] = -[\vec{z}, \vec{y}, \vec{x}] = [\vec{y}, \vec{z}, \vec{x}] = -[\vec{y}, \vec{x}, \vec{z}]$$

- El producto mixto $[\vec{x}, \vec{y}, \vec{z}]$ es nulo si y sólo si el sistema de vectores $(\vec{x}, \vec{y}, \vec{z})$ es linealmente dependiente.

DEMOSTRACION

Estas tres propiedades son consecuencias inmediatas de las propiedades de los determinantes: basta aplicar éstas (véase [034] y [040]) a las columnas de los determinantes que definen los productos escalares que se citan en el enunciado.

EJERCICIO

Recurriendo a que, según se sabe de geometría elemental, el volumen de un paralelepípedo es igual al producto del área de su base por su altura, compruébese que:

[156]

En el espacio vectorial euclídeo canónico $\mathbb{R}^3$, el volumen del paralelepípedo que tiene por aristas a los vectores $\bar{x}, \bar{y}, \bar{z} \in \mathbb{R}^3$ es igual a:

$$\text{Volumen} = \varepsilon[\bar{x}, \bar{y}, \bar{z}],$$

donde

$$\varepsilon = \begin{cases} 1 & \text{si } (\bar{x}, \bar{y}, \bar{z}) \text{ tiene orientación positiva} \\ -1 & \text{si } (\bar{x}, \bar{y}, \bar{z}) \text{ tiene orientación negativa} \end{cases}$$

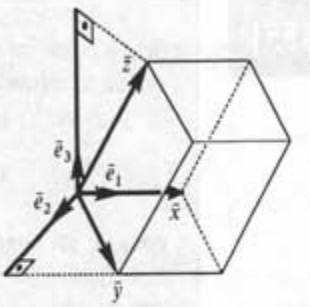


Figura 2.

RESOLUCION

Sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ la siguiente base ortonormal de $\mathbb{R}^3$: $\bar{e}_1$ es unitario y con la dirección y sentido de $\bar{x}$; $\bar{e}_2$ es unitario, está situado en el plano de $\bar{x}$ e $\bar{y}$, es ortogonal a $\bar{e}_1$, y tal que $(\bar{e}_1, \bar{e}_2)$ tiene la misma orientación que $(\bar{x}, \bar{y})$; $\bar{e}_3$ es unitario, es ortogonal a $\bar{e}_1$ y a $\bar{e}_2$ y es tal que $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ tiene la misma orientación que $(\bar{x}, \bar{y}, \bar{z})$. Las coordenadas de $\bar{x}$, $\bar{y}$ y $\bar{z}$ en la base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ son del tipo

$$\bar{x}(a, 0, 0) \quad , \quad \bar{y}(a', b, 0) \quad , \quad z(a'', b', c)$$

donde a y b son la «base» y la «altura» del paralelogramo de aristas $\bar{x}$ e $\bar{y}$, que se toma como base del paralelepípedo dado, el cual tiene entonces por «altura» al número c. Nótese que el volumen del paralelepípedo es igual a

$$\text{Volumen} = abc \qquad [1]$$

La base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ es directa (inversa) si $(\bar{x}, \bar{y}, \bar{z})$ tiene orientación positiva (negativa) esto es, si $\varepsilon = 1$ $(\varepsilon = -1)$. Por tanto, la base $(\bar{e}_1, \bar{e}_2, \varepsilon\bar{e}_3)$ es directa. Expresando en ella el producto mixto $[\bar{x}, \bar{y}, \bar{z}]$, se tiene:

$$[\bar{x}, \bar{y}, \bar{z}] = \begin{vmatrix} a & a' & \varepsilon a'' \\ 0 & b & \varepsilon b' \\ 0 & 0 & \varepsilon c \end{vmatrix} = \varepsilon abc \qquad [2]$$

De [1] y [2] y como $1/\varepsilon = \varepsilon$, se desprende la igualdad a demostrar.

OBSERVACIONES

1. En un espacio euclídeo V_n, de dimensión n, orientado, se llama producto mixto de un sistema $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ de n vectores de V_n al valor del determinante, de orden n, cuyas columnas son las de coordenadas de $\bar{x}_1$, $\bar{x}_2$, ..., $\bar{x}_n$, en este orden, respecto de una base ortonormal directa cualquiera de V_n; este valor se denotará poniendo $[\bar{x}_1, \bar{x}_2, ..., \bar{x}_n]$. Esta definición es consistente ya que el valor del referido determinante no depende de la base (ortonormal directa) que se considere, lo que se prueba de igual modo que en el caso $n = 3$.
   En V_n, el producto mixto sigue siendo una forma lineal respecto de cada una de las variables (es decir, n-lineal) y antisimétrica.
2. En un espacio vectorial euclídeo V_n, de dimensión n, orientado, se llama «paralelotopo» que tiene por aristas a los vectores $\bar{x}_1, \bar{x}_2, ..., \bar{x}_n$ al conjunto que forman los siguientes vectores $\bar{x}$:

$$\bar{x} = \lambda_1\bar{x}_1 + \lambda_2\bar{x}_2 + \cdots + \lambda_n\bar{x}_n \quad , \quad \text{para } 0 \leq \lambda_i \leq 1 \ (i = 1, 2, ..., n)$$

   Se define el «volumen orientado» de dicho paralelotopo como el producto mixto de sus n aristas. El signo del volumen depende del orden en el que se consideren sus aristas; el volumen es positivo si las aristas se toman en orden directo, o sea, positivamente orientadas.

8.9. PRODUCTO VECTORIAL

Dados dos vectores $\bar{x}_1$ y $\bar{x}_2$, de un espacio vectorial euclídeo V_3, orientado, de dimensión 3, para definir su producto vectorial vamos a recurrir a que existe cierto vector $\bar{p}$ (dependiente de $\bar{x}_1$ y $\bar{x}_2$) tal que, para cualquiera que sea $\bar{x} \in V_3$, se verifica que $[\bar{x}_1, \bar{x}_2, \bar{x}] = \bar{p} \cdot \bar{x}$. Para comprobar esto, necesitamos del siguiente resultado previo:

PROPOSICION

[157] Si V_n es un espacio vectorial euclídeo, de dimensión finita n, y dada una forma lineal $f : V_n \to \mathbb{R}$, existe un único vector $\bar{p} \in V_n$ tal que $f(\bar{x}) = \bar{p} \cdot \bar{x}$ para todo $\bar{x} \in V_n$.

DEMOSTRACION

Si f es nula ($f = o$), entonces $\bar{p} = \bar{o}$ es, obviamente, solución y es solución única pues, para $\bar{p} \neq \bar{o}$, $\bar{p} \cdot \bar{p} \neq 0$ luego $\bar{p} \cdot \bar{x}$ no es nulo para todo $\bar{x} \in V_n$. Supongamos ahora que es $f \neq o$. Como la imagen de f tiene dimensión 1, resulta que su núcleo $U = \text{Nuc}(f)$ es un subespacio de dimensión $n - 1$. El subespacio U tiene, pues, subespacio suplementario ortogonal de dimensión 1; sea $\bar{e}$ vector (no nulo) de la dirección de este subespacio. Si $\bar{x} \in U$ es $f(\bar{x}) = 0$,

luego $\bar{p}$ ha de ser tal que $\bar{p} \cdot \bar{x} = 0$ para todo $\bar{x} \in U$, es decir, $\bar{p}$ tiene que ser ortogonal a U, esto es, debe ser $\bar{p} = h\bar{e}$ para algún $h \in \mathbb{R}$, que vamos a determinar. Como $U \oplus \mathcal{V}(\bar{e}) = V_n$, cualquier $\bar{x} \in V_n$ se puede expresar de manera única como suma del tipo $\bar{x} = \lambda\bar{e} + \bar{u}$, con $\lambda \in \mathbb{R}$ y $u \in U$, y por tanto:

$$f(\bar{x}) = f(\lambda\bar{e} + \bar{u}) = \lambda f(\bar{e}) + f(\bar{u}) = \lambda f(\bar{e}) + 0 = \lambda f(\bar{e})$$

$$\bar{p} \cdot \bar{x} = (h\bar{e}) \cdot (\lambda\bar{e} + \bar{u}) = h\lambda\|\bar{e}\|^2 + h\,\bar{e} \cdot \bar{u} = h\lambda + 0 = h\lambda$$

En consecuencia, la relación $f(\bar{x}) = \bar{p} \cdot \bar{x}$ se verifica, para todo $\bar{x} \in V$, si y sólo si se toma $\bar{p} = h\bar{e}$ con $h = f(\bar{e})$.

PRODUCTO VECTORIAL

[158]

Sea V_3 un espacio vectorial euclídeo, orientado, de dimensión 3.

I) Para cualesquiera que sean los vectores $\bar{u}, \bar{v} \in V$ existe un único vector $\bar{p} \in V_3$ tal que:

$$[\bar{u}, \bar{v}, \bar{x}] = \bar{p} \cdot \bar{x} \quad \text{para todo } \bar{x} \in V$$

A este vector $\bar{p}$ se le llama *producto vectorial* de $\bar{u}$ por $\bar{v}$ y se pone:

$$\bar{p} = \bar{u} \wedge \bar{v}$$

De acuerdo con esta definición, para cualesquiera que sean los vectores $\bar{u}, \bar{v}, \bar{w} \in V_3$, se verifica la siguiente relación:

$$(\bar{u} \wedge \bar{v}) \cdot \bar{w} = [\bar{u}, \bar{v}, \bar{w}]$$

II) El producto vectorial $\bar{u} \wedge \bar{v}$ (donde $\bar{u}, \bar{v} \in V_3$) se puede también definir como el vector de V_3 tal que: si $\bar{u}$ y $\bar{v}$ son linealmente dependientes, entonces $\bar{u} \wedge \bar{v} = \bar{o}$; si $\bar{u}$ y $\bar{v}$ son independientes, entonces

1. $\bar{u} \wedge \bar{v}$ es ortogonal a $\bar{u}$ y a $\bar{v}$.
2. La orientación de $(\bar{u}, \bar{v}, \bar{u} \wedge \bar{v})$ es positiva.
3. $\|\bar{u} \wedge \bar{v}\| = \|\bar{u}\|\,\|\bar{v}\|\,\text{sen}\,(\bar{u}, \bar{v})$.

A partir de esta segunda forma de definir el producto vectorial, se puede también definir el producto mixto recurriendo a la relación:

$$[\bar{u}, \bar{v}, \bar{w}] = (\bar{u} \wedge \bar{v}) \cdot \bar{w}$$

DEMOSTRACION

I) Según se dijo al hablar del producto mixto en [155], la aplicación

$$f : V_3 \to \mathbb{R}, \quad \bar{x} \mapsto f(\bar{x}) = [\bar{u}, \bar{v}, \bar{x}]$$

es una forma lineal. De acuerdo, entonces, con la proposición [157], existe un único vector $\bar{p} \in V_3$ tal ~~que~~ $f(\bar{x}) = \bar{p} \cdot \bar{x}$, es decir, $[\bar{u}, \bar{v}, \bar{x}] = \bar{p} \cdot \bar{x}$, para todo $\bar{x} \in V_3$.

Nótese que, de acuerdo con esta definición de $\bar{u} \wedge \bar{v}$, los vectores $\bar{u}$ y $\bar{v}$ son linealmente dependientes si y sólo si $\bar{u} \wedge \bar{v} = \bar{o}$. En efecto: el vector $\bar{p} = \bar{u} \wedge \bar{v}$ será nulo si y sólo si $\bar{p} \cdot \bar{x} = 0$ para todo $\bar{x} \in V_3$, es decir, si $[\bar{u}, \bar{v}, \bar{x}] = 0$ para todo $\bar{x} \in V_3$, lo que equivale a que $\bar{u}$ y $\bar{v}$ sean linealmente dependientes.

II) Supongamos en primer lugar que $\bar{u}$ y $\bar{v}$ son linealmente dependientes, o sea, que $\operatorname{sen}(\bar{u}, \bar{v}) = 0$; esto equivale (por la condición 3.ª) a que según la definición II sea $\|\bar{u} \wedge \bar{v}\| = 0$, esto es. $\bar{u} \wedge \bar{v} = \bar{o}$. Resulta, pues, que las definiciones I y II coinciden si $\bar{u}$ y $\bar{v}$ son linealmente dependientes y que, según ambas, $\bar{u} \wedge \bar{v} = \bar{o}$ equivale a que $\bar{u}$ y $\bar{v}$ sean linealmente dependientes. Supondremos, pues, a partir de aquí, que $\bar{u}$ y $\bar{v}$ son independientes (o sea, $\bar{u} \wedge \bar{v} \neq \bar{o}$):

Como el $\bar{u} \wedge \bar{v}$ definido en I) existe y es único y, obviamente, lo mismo ocurre con el $\bar{u} \wedge \bar{v}$ de II), para comprobar que ambos modos de introducir $\bar{u} \wedge \bar{v}$ son equivalentes, bastará con mostrar que con la definición I) se verifican las propiedades que se utilizan para la definición II).

1. El vector $\bar{u} \wedge \bar{v}$ es ortogonal a $\bar{u}$ y a $\bar{v}$ ya que

$$(\bar{u} \wedge \bar{v}) \cdot \bar{u} = [\bar{u}, \bar{v}, \bar{u}] = 0 \quad \text{y} \quad (\bar{u} \wedge \bar{v}) \cdot \bar{v} = [\bar{u}, \bar{v}, \bar{v}] = 0$$

2. Hay que comprobar que el producto mixto $[\bar{u}, \bar{v}, \bar{u} \wedge \bar{v}]$ es positivo. Así ocurre, ya que, como $\bar{u} \wedge \bar{v} \neq \bar{o}$, es $\|\bar{u} \wedge \bar{v}\| > 0$ y entonces:

$$0 < \|\bar{u} \wedge \bar{v}\|^2 = (\bar{u} \wedge \bar{v}) \cdot (\bar{u} \wedge \bar{v}) = [\bar{u}, \bar{v}, \bar{u} \wedge \bar{v}]$$

3. Sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ la siguiente base ortonormal directa: $\bar{e}_1$ es unitario de la dirección y sentido de $\bar{u}$; $\bar{e}_2$ es unitario, ortogonal a $\bar{e}_1$, situado en el plano de $\bar{u}$ y $\bar{v}$ y tal que $(\bar{e}_1, \bar{e}_2)$ está igual orientado que $(\bar{u}, \bar{v})$; $\bar{e}_3$ tiene la dirección y el sentido de $\bar{u} \wedge \bar{v}$, con lo que la orientación de $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ es directa. En esta base, las coordenadas de $\bar{u}$, $\bar{v}$ y $\bar{u} \wedge \bar{v}$ son

$$\bar{u}(\|\bar{u}\|, 0, 0) \quad , \quad \bar{v}(\|\bar{v}\| \cos(\bar{u}, \bar{v}), \|\bar{v}\| \operatorname{sen}(\bar{u}, \bar{v}), 0)$$

y

$$\bar{u} \wedge \bar{v}(0, 0, \|\bar{u} \wedge \bar{v}\|)$$

Por tanto, recurriendo a la expresión en coordenadas del producto mixto y recurriendo a la última igualdad, se obtiene:

$$\|\bar{u} \wedge \bar{v}\|^2 = [\bar{u}, \bar{v}, \bar{u} \wedge \bar{v}] = \begin{vmatrix} \|\bar{u}\| & \|\bar{v}\| \cos(\bar{u}, \bar{v}) & 0 \\ 0 & \|\bar{v}\| \operatorname{sen}(\bar{u}, \bar{v}) & 0 \\ 0 & 0 & \|\bar{u} \wedge \bar{v}\| \end{vmatrix} =$$

$$= \|\bar{u}\| \, \|\bar{v}\| \operatorname{sen}(\bar{u}, \bar{v}) \|\bar{u} \wedge \bar{v}\|$$

Dividiendo por $\|\bar{u} \wedge \bar{v}\|$ (que no es nulo) se obtiene la expresión del enunciado.

PROPIEDADES

[159]

El producto vectorial, entre vectores de un espacio vectorial euclídeo V_3, de dimensión 3, orientado, goza (además de las anteriores) de las siguientes propiedades (en las que $\bar{u}, \bar{u}', \bar{v}, \bar{v}' \in V_3$ y $\lambda \in \mathbb{R}$ son cualesquiera):

1. $\bar{u} \wedge \bar{v} = \bar{o}$ si y sólo si $\bar{u}$ y $\bar{v}$ son linealmente dependientes.
2. $\bar{u} \wedge (\bar{v} + \bar{v}') = \bar{u} \wedge \bar{v} + \bar{u} \wedge \bar{v}'$
3. $(\bar{u} + \bar{u}') \wedge \bar{v} = \bar{u} \wedge \bar{v} + \bar{u}' \wedge \bar{v}$
4. $(\lambda \bar{u}) \wedge \bar{v} = \bar{u} \wedge (\lambda \bar{v}) = \lambda(\bar{u} \wedge \bar{v})$
5. $\bar{u} \wedge \bar{v} = -\bar{v} \wedge \bar{u}$

DEMOSTRACION

1. Aun cuando esta propiedad ya se comprobó a lo largo de la demostración de [158], no viene mal considerarla de nuevo. Si $\bar{u}$ y $\bar{v}$ son linealmente dependientes, entonces $[\bar{u}, \bar{v}, \bar{x}] = 0$ para todo $\bar{x} \in V_3$, o sea $(\bar{u} \wedge \bar{v}) \cdot \bar{x} = 0$ para todo $x \in V_3$, luego $\bar{u} \wedge \bar{v} = \bar{o}$. Si $\bar{u}$ y $\bar{v}$ son linealmente independientes, se puede elegir $\bar{x}_0 \in V_3$ de manera que el sistema $(\bar{u}, \bar{v}, \bar{x}_0)$ sea independiente, de manera que $[\bar{u}, \bar{v}, \bar{x}_0] \neq 0$, o sea $(\bar{u} \wedge \bar{v}) \cdot \bar{x}_0 \neq 0$, luego $\bar{u} \wedge \bar{v} \neq \bar{o}$.

2. $[\bar{u}, \bar{v} + \bar{v}', \bar{x}] = [\bar{u}, \bar{v}, \bar{x}] + [\bar{u}, \bar{v}', x] \quad \forall \bar{x} \in V$, luego

$[\bar{u} \wedge (\bar{v} + \bar{v}')] \cdot \bar{x} = (\bar{u} \wedge \bar{v}) \cdot \bar{x} + (\bar{u} \wedge \bar{v}') \cdot \bar{x} \quad \forall \bar{x} \in V$, luego

$[\bar{u} \wedge (\bar{v} + \bar{v}')] \cdot \bar{x} = [(\bar{u} \wedge \bar{v}) + (\bar{u} \wedge \bar{v}')] \cdot \bar{x} \quad \forall \bar{x} \in V$, luego

$$\bar{u} \wedge (\bar{v} + \bar{v}') = (\bar{u} \wedge \bar{v}) + (u \wedge \bar{v}')$$

3. y 4. Se prueban como la propiedad anterior.

5. Recurramos a que, según ya sabemos, es:

$$\bar{w} \wedge \bar{w} = \bar{o}, \quad \forall \bar{w} \in V_3$$

Tomando $\bar{w} = \bar{u} + \bar{v}$, se tiene:

$$\bar{o} = (\bar{u} + \bar{v}) \wedge (\bar{u} + \bar{v}) = \bar{u} \wedge \bar{u} + \bar{u} \wedge \bar{v} + \bar{v} \wedge \bar{u} + \bar{v} \wedge \bar{v} =$$

$$= \bar{o} + \bar{u} \wedge \bar{v} + \bar{v} \wedge \bar{u} + \bar{o}$$

así que

$$\bar{u} \wedge \bar{v} + \bar{v} \wedge \bar{u} = \bar{o}, \quad \text{o sea,} \quad \bar{u} \wedge \bar{v} = -\bar{v} \wedge \bar{u}$$

EXPRESION DEL PRODUCTO VECTORIAL

[160]

Sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ una base ortonormal directa del espacio vectorial euclídeo V_3, de dimensión 3 y orientado. Se verifica que

1. $\bar{e}_1 \wedge \bar{e}_2 = \bar{e}_3 \quad , \quad \bar{e}_2 \wedge \bar{e}_3 = \bar{e}_1 \quad , \quad \bar{e}_3 \wedge \bar{e}_1 = \bar{e}_2$.
2. Si las coordenadas, en dicha base, de dos vectores $\bar{u}, \bar{v} \in V_3$ son

$$\bar{u}(u_1, u_2, u_3) \quad \text{y} \quad \bar{v}(v_1, v_2, v_3)$$

entonces

$$\bar{u} \wedge \bar{v} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \bar{e}_1 + \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix} \bar{e}_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \bar{e}_3 = \begin{vmatrix} \bar{e}_1 & \bar{e}_2 & \bar{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \quad (*)$$

(*) Este último miembro es una forma simbólica de representar el segundo miembro. Esta expresión, que se recuerda con facilidad, no es un determinante, pues su primera fila no está formada por escalares sino por vectores. No obstante, le es de aplicación de fórmula del desarrollo de un determinante por elemento de la primera fila.

DEMOSTRACION

1. Como $\bar{e}_1 \wedge \bar{e}_2$ es ortogonal a $\bar{e}_1$ y a $\bar{e}_2$, resulta que $\bar{e}_1 \wedge \bar{e}_2 = \lambda \bar{e}_3$ para cierto $\lambda \in \mathbb{R}$; como $(\bar{e}_1\, \bar{e}_2, \bar{e}_1 \wedge \bar{e}_2)$ tiene orientación positiva, debe ser $\lambda > 0$; como áng $(\bar{e}_1, \bar{e}_2) = \pi/2$, será

$$|\lambda| = \|\lambda \bar{e}_3\| = \|\bar{e}_1 \wedge \bar{e}_2\| = \|\bar{e}_1\| \, \|\bar{e}_2\| \operatorname{sen}(\bar{e}_1, \bar{e}_2) = 1$$

Por tanto, es $\lambda = 1$, de donde resulta que $\bar{e}_1 \wedge \bar{e}_2 = \bar{e}_3$.
Análogamente se comprueba que $\bar{e}_2 \wedge \bar{e}_3 = \bar{e}$ y que $\bar{e}_3 \wedge \bar{e}_1 = \bar{e}_2$.

2. Recurriendo a las propiedades anteriormente demostradas, se obtiene:

$$\bar{u} \wedge \bar{v} = \left(\sum_{i=1}^{3} u_i \bar{e}_i \right) \wedge \left(\sum_{i=1}^{3} v_j \bar{e}_j \right) = \sum_{i,j=1}^{3} u_i v_j (\bar{e}_i \wedge \bar{e}_j)$$

teniendo en cuenta las tres igualdades precedentes, y como $\bar{e}_i \wedge \bar{e}_j = -\bar{e}_j \wedge e_i$ y como $\bar{e}_i \wedge \bar{e}_i = \bar{o}$, resulta que:

$$\bar{u} \wedge \bar{v} = (u_2 v_3 - v_3 v_2)\bar{e}_1 + (u_3 v_1 - u_1 v_3)\bar{e}_2 + (u_1 v_2 - u_2 v_1)\bar{e}_3$$

que es lo que se quería comprobar.

OBSERVACION

El producto vectorial puede generalizarse, a un espacio vectorial euclídeo de dimensión finita n cualquiera, V_n, orientado. Para ello, se recurre al producto mixto de n vectores de V_n.

Procediendo del mismo modo que en el caso $n = 3$, se prueba que: para cualesquiera que sean los vectores $\bar{u}_1, \bar{u}_2, ..., \bar{u}_{n-1}$ de V_n, existe un único vector $\bar{p} \in V_n$ tal que

$$[\bar{u}_1, \bar{u}_2, ..., \bar{u}_{n-1}, \bar{x}] = \bar{p} \cdot \bar{x}, \quad \forall \bar{x} \in V_n$$

A este vector se le llama producto vectorial de los vectores $\bar{u}_1, \bar{u}_2, ..., \bar{u}_{n-1}$:

$$\bar{p} = \bar{u}_1 \wedge \bar{u}_2 \wedge \cdots \wedge \bar{u}_{n-1}$$

El producto vectorial es lineal respecto de cada una de sus $n - 1$ variables. El producto vectorial es nulo si y sólo si los vectores $\bar{u}_1, \bar{u}_2, ..., \bar{u}_{n-1}$ son linealmente dependientes. En una base ortonormal directa, las coordenadas de $\bar{p}$ son los adjuntos de los elementos de la primera fila de un determinante cuyas demás filas son las formadas por las coordenadas de los vectores $\bar{u}_1, \bar{u}_2, ..., \bar{u}_{n-1}$ en la referida base.

EJERCICIO

Compruébese que, para cualesquiera vectores $\bar{u}$, $\bar{v}$, de un espacio vectorial euclídeo V_3, de dimensión 3 y orientado, se verifica:

$$\bar{u} \wedge (\bar{u} \wedge \bar{v}) = (\bar{u} \cdot \bar{v})\bar{u} - (\bar{u} \cdot \bar{u})\bar{v}$$

RESOLUCION

Descompongamos $\bar{v} = \bar{v}_1 + \bar{v}_2$, con $\bar{v}_1$ en la dirección de $\bar{u}$ y $\bar{v}_2$ ortogonal a $\bar{u}$. Con ello, como $\bar{u} \wedge \bar{v}_1 = \bar{o}$ y $\bar{u} \cdot \bar{v} = \bar{u} \cdot \bar{v}_1$, la relación a demostrar toma la forma

$$\bar{u} \wedge (\bar{u} \wedge \bar{v}_2) = (\bar{u} \cdot \bar{v}_1)\bar{u} - (\bar{u} \cdot \bar{u})(\bar{v}_1 + \bar{v}_2)$$

Si son $\bar{e}_1$ y $\bar{e}_2$ los vectores unitarios de las direcciones y sentidos de $\bar{u}$ y de $\bar{v}_2$, se puede poner:

$$\bar{u} \wedge (\bar{u} \wedge \bar{v}_2) = -\|\bar{u}\|^2 \|\bar{v}_2\| \bar{e}_2$$

$$(\bar{u} \cdot \bar{v}_1)\bar{u} - (\bar{u} \cdot \bar{u})(\bar{v}_1 + \bar{v}_2) = \|\bar{u}\|^2 \|\bar{v}_1\| \bar{e}_1 - \|\bar{u}\|^2 \|\bar{v}_1\| \bar{e}_1 - \|\bar{u}\|^2 \|\bar{v}_2\| \bar{e}_2$$

luego ambos vectores son iguales, como había que comprobar.

Ejercicios y problemas a la parte III

ENUNCIADOS

III.1. Sea $f: V \times V \to \mathbb{R}$ una forma bilineal, en un espacio vectorial real V. Compruébese que f es alternada si, y sólo si, f es antisimétrica, sabiendo que:

- f se dice alternada si $f(\bar{x}, \bar{x}) = 0$ para todo $\bar{x} \in V$.
- f se dice antisimétrica si $f(\bar{x}, \bar{y}) = -f(\bar{y}, \bar{x})$ para cualesquiera $\bar{x}, \bar{y} \in V$.

III.2. Sea $f: V \times V \to K$ una forma bilineal, en el espacio vectorial V de dimensión n sobre el cuerpo K. Sea $A = [a_{ij}]$ la matriz de f en una base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V. Determinar las transformaciones que se producen en la matriz A cuando en la base dada se realizan las siguientes manipulaciones:

1. Multiplicar $\bar{e}_1$ por $\lambda \neq 0$.
2. Permutar entre sí los vectores $\bar{e}_1$ y $\bar{e}_2$.
3. Sumar a $\bar{e}_1$ el vector $\lambda \bar{e}_2$.

III.3. Sea $f: V \times V \to K$ una forma bilineal, en el espacio vectorial V sobre un cuerpo K. Sea $C \subset V$ un conjunto cualquiera de vectores y considérense los nuevos conjuntos C_1 y C_2 siguientes:

$$C_1 = \{\bar{u} \in V / f(\bar{x}, \bar{u}) = 0, \forall \bar{x} \in C\}$$

y

$$C_2 = \{\bar{u} \in V / f(\bar{u}, \bar{x}) = 0, \forall \bar{x} \in C\}$$

1. Pruébese que C_1 y C_2 son subespacios de V.
2. Si $D \subset V$ es tal que $C \subset D$, hallar las relaciones de inclusión que hay entre C_1 y D_1 y entre C_2 y D_2.
3. Hallar C_1 y C_2 si $C = O$.

III.4. Sea V un espacio vectorial de dimensión finita sobre un cuerpo K y sea $f: V \times V \to K$ una forma bilineal simétrica ordinaria (no degenerada). Dada una aplicación lineal $\varphi: V \to V$, pruébese que existe una aplicación lineal $\varphi^*: V \to V$ tal que $f(\varphi(\bar{x}), \bar{y}) = f(\bar{x}, \varphi^*(y))$ para cualesquiera $\bar{x}, \bar{y} \in V$.

Si $\psi: V \to V$ es otro endomorfismo, hallar (en función de φ^* y de ψ^*):

$$(\varphi \circ \psi)^* \quad (\lambda\varphi + \mu\psi)^* \quad (\varphi^*)^*$$

III.5. Sea $\omega: V \to K$ una forma cuadrática, en un espacio vectorial V sobre el cuerpo K. Si f es la forma polar de ω, compruébese que para cualesquiera vectores $\bar{u}, \bar{v} \in V$ se verifica que:

$$4f(\bar{u}, \bar{v}) = \omega(\bar{u} + \bar{v}) - \omega(\bar{u} - \bar{v})$$

$$2[\omega(\bar{u}) + \omega(\bar{v})] = \omega(\bar{u} + \bar{v}) + \omega(\bar{u} - \bar{v})$$

III.6. Sea V el espacio vectorial de las funciones continuas de $\mathbb{R}$ en $\mathbb{R}$ y considérese la aplicación $\omega: V \to \mathbb{R}$, dada por:

$$\omega(\varphi) = \int_{-a}^{a} \varphi(x + b)\varphi(x - b)\, dx$$

($a, b \in \mathbb{R}$ dados). Se pide:

1. Comprobar que ω es una forma cuadrática, hallando su forma polar.
2. Si V fuese el espacio vectorial de los polinomios de grado menor o igual que dos, hallar la matriz A de ω en la base $(1, x, x^2)$.
3. En el supuesto del apartado anterior y si $a = \sqrt{3}$ y $b = 1$, hallar el núcleo de ω y diagonalizarla.

III.7. Sea V un espacio vectorial sobre un cuerpo K; sea $\varphi: V \to V$ una aplicación lineal; sea $\omega: V \to K$ una forma cuadrática. Comprobar que $\omega \circ \varphi$ es una forma cuadrática. Si V tiene dimensión finita, hállese la matriz de $\omega \circ \varphi$, respecto de una cierta base de V, en función de las matrices M y A de φ y ω, en dicha base. ¿Qué hay que exigir a φ y ω para que $\omega \circ \varphi$ sea regular?

III.8. Sea $\omega: V \to K$ una forma cuadrática, en un espacio vectorial V sobre el cuerpo K. Sea $\bar{u} \in V$ un vector que no pertenece al núcleo de ω. Demuéstrese que existe algún vector $\bar{v} \in V$ tal que el subespacio $\mathscr{C}(\bar{u})$, conjugado de $\bar{u}$, es suplementario del $\mathscr{V}(\bar{v})$, engendrado por $\bar{v}$.

III.9. Sea $\omega: V \to K$ una forma cuadrática, en un espacio vectorial V sobre el cuerpo K; sea $D \subset V$ un conjunto cualquiera de vectores. Pruébese que el conjunto D^c, que forman los vectores que son conjugados de todos los vectores de D, es un

subespacio vectorial (que se llama subespacio conjugado de D).

III.10. Sea $\omega: K^n \to K$ una forma cuadrática no nula. Compruébese que $\omega(x_1, x_2, ..., x_n)$ se puede expresar como producto de dos expresiones lineales si, y sólo si, su matriz (en base canónica) tiene rango unidad.

III.11. Sea V el espacio vectorial real de las matrices cuadradas de tamaño $n \times n$. Considérese la aplicación $f: V \times V \to \mathbb{R}$ siguiente:

$$(M, N) \mapsto f(M, N) = \text{traza } (MN) - \\ - (\text{traza } M)\ (\text{traza } N)$$

1. Probar que f es una forma bilineal simétrica.
2. Para $n = 2$, hallar la forma cuadrática $\omega: V \to \mathbb{R}$ asociada a f, determinando la expresión de $\omega(M)$ para

$$M = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$$

3. Para $n = 2$, hallar la matriz A asociada a ω en la base cuyos elementos son:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ y } \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

4. Para $n = 2$, hallar el subespacio de V formado por las matrices conjugadas, respecto de ω, de todas las matrices antisimétricas de V.

III.12. Sea V el espacio vectorial de los polinomios reales $p(x) = a_0 + a_1 x + a_2 x^2$ de grado menor o igual que dos y considérese la aplicación $\omega: V \to \mathbb{R}$, $\omega(p(x)) = p(\alpha)p(\beta)$, donde $\alpha, \beta \in \mathbb{R}$ son dados. Se pide:

1. Probar que ω es una forma cuadrática, hallando su forma polar.
2. Hallar la matriza A en la base $(1, x, x^2)$ de V.
3. Relación entre α y β para que 1 y x sean conjugados respecto de ω.
4. En el supuesto del apartado anterior, hallar el núcleo de ω y diagonalizarla.

III.13. Sea $\omega: V \to \mathbb{C}$ una forma cuadrática, en el espacio vectorial complejo V de dimensión finita. Pruébese que existe alguna base de V en la que la matriz $D = [d_{ij}]$ de ω es diagonal y tiene

$$d_{ii} = \begin{cases} 1, & \text{si } i \leq r \\ 0, & \text{si } i > r \end{cases} \quad (r = \text{rang } \omega)$$

III.14. De una forma cuadrática $\omega: \mathbb{R}^3 \to \mathbb{R}$ se sabe que:

- Los vectores $(0, 1, 0)$ y $(0, 1, -1)$ son conjugados respecto de ω.
- $(-1, 0, 1)$ es un vector del núcleo de ω.
- $\omega(1, 0, 0) = 1$.
- La traza de la matriz A de ω en la base canónica vale 0.

1. Hallar la matriz A.
2. Hallar la expresión canónica de ω obteniendo, también, la base en la que se consigue.

III.15. Se considera la forma cuadrática $\omega: \mathbb{R}^3 \to \mathbb{R}$ que tiene por expresión (en la base canónica) a:

$$\omega(x, y, z) = \alpha x^2 + (\alpha + 3)y^2 + (\alpha + 2)z^2 + \\ + 2(\alpha + 1)xy + 2xz + 4yz$$

1. Diagonalizar ω, determinando la base de la diagonalización que se obtenga, para los distintos valores de $\alpha \in \mathbb{R}$.
2. Hallar el núcleo de ω, en función de α.
3. Para $\alpha = -1$, comprobar que los vectores autoconjugados respecto de ω forman dos subespacios vectoriales de $\mathbb{R}^3$.
4. Para $\alpha = -1$, expresar ω en una base formada por vectores autoconjugados respecto de ω.

III.16. Sea $\varphi: \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ la forma bilineal que, respecto de la base canónica, tiene asociada la matriz

$$\begin{bmatrix} 1 & 2 & -2 & -3 \\ 0 & 1 & 2 & -2 \\ 2 & 0 & 1 & 2 \\ 3 & 2 & 0 & \alpha \end{bmatrix} \quad (\alpha \in \mathbb{R} \text{ dado})$$

1. Compruébese, hallando su matriz A (simétrica), que $\omega: \mathbb{R}^4 \to \mathbb{R}$, $\omega(\bar{x}) = \varphi(\bar{x}, \bar{x})$ es una forma cuadrática.
2. Hallar el núcleo de ω, en función de α.
3. Diagonalizar ω, obteniendo la base de la diagonalización que se de.
4. Para $\alpha = 4$, hallar una base de $\mathbb{R}^4$ en la que la matriz de ω tiene nulos todos los elementos de su diagonal; hallar también la expresión de $\omega(\bar{x})$ en esta base.

III.17. Se dice que una forma cuadrática real $\omega: V \to \mathbb{R}$ (donde V es un espacio vectorial real cualquiera)

es definida si para todo vector $\bar{x} \neq \bar{0}$ de V es $\omega(\bar{x}) \neq 0$. Demuéstrese que si ω es definida, entonces ω es definida positiva o es definida negativa.

III.18. Sea M una matriz cuadrada regular. Pruébese que M^tM es una matriz simétrica definida positiva.

III.19. Sea A una matriz cuadrada $n \times n$. Hallar una condición necesaria y suficiente para que A se pueda poner en la forma $A = MM^t$ para cierta matriz regular M.

III.20. Sea $\omega: \mathbb{R}^n \to \mathbb{R}$ la forma cuadrática no nula que, en base canónica tiene la expresión:

$$\omega(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

con $a_{11} \neq 0$ ($a_{ij} = a_{ji}$). Hallar un cambio de coordenadas $(x_1, x_2, \ldots, x_n) \rightarrowtail (y_1, y_2, \ldots, y_n)$ de manera que

$$\omega(y_1, y_2, \ldots, y_n) = a_{11}y_1^2 + \omega_1(y_2, \ldots, y_n)$$

con ω_1 cuadrática. Si fuese $a_{11} = 0$, dese un cambio de coordenadas de modo que el nuevo elemento de lugar 11 no sea nulo.

III.21. Sea $\omega: V \to \mathbb{R}$ una forma cuadrática, en el espacio vectorial real V, y sea f la forma polar de ω. Demuéstrese que si ω es definida positiva, entonces para cualesquiera $\bar{u}, \bar{v} \in V$ si verifica que:

$$f(\bar{u}, \bar{v})^2 \leq \omega(\bar{u})\omega(\bar{v})$$

(desigualdad de Schwarz).

III.22. Sea V el espacio vectorial de las matrices cuadradas de tamaño 2×2 y considérese en él la base usual, esto es, la formada por:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ y } \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Compruébese que, para cualquiera que sea $c \in \mathbb{R}$, la aplicación $f: V \times V \to \mathbb{R}$ dado por:

$$f(M, N) = c\ (\text{traza } MN) - (\text{traza } M)(\text{traza } N)$$

es bilineal simétrica y:

1. Hallar la matriz A de f en la base dada.
2. Diagonalizar ω, según los valores de c, obteniendo la base en la que ω adopte la forma diagonal que se obtenga.
3. Estudiar, en función de c, si ω es definida o semidefinida.

III.23. Sea $\omega: \mathbb{R}^4 \to \mathbb{R}$ la forma cuadrática que, en la base canónica, tiene por expresión a:

$$\omega(x_1, x_2, x_3, x_4) = 2x_1^2 + \alpha x_2^2 + 2x_3^2 + \alpha x_4^2 + 2x_1x_3 + 2x_2x_4$$

Se pide, en función del valor que tome el parámetro $\alpha \in \mathbb{R}$:

1. Diagonalizar ω.
2. El rango y la signatura de ω.
3. Estudiar si ω es definida o semidefinida.

III.24. Determínense los valores $\alpha \in \mathbb{R}$ para los que

$$(x, y, z) \cdot (x', y', z') = (2x - y + z)(2x' - y' + z') + \alpha(xx' + yy' + zz')$$

define un producto escalar en $\mathbb{R}^3$. Para $\alpha = 1$, hallar una base ortonormal de $\mathbb{R}^3$ (respecto del anterior producto escalar).

III.25. Hallar la relación que han de verificar los números reales α y β para que

$$\bar{x} \cdot \bar{y} = 2^{\alpha}x_1y_1 + 2^{\alpha\beta}(x_1y_2 + x_2y_1) + 2^{\beta^2}x_2y_2$$

donde

$$\bar{x} = (x_1, x_2) \quad \text{e} \quad \bar{y} = (y_1, y_2)$$

determine un producto escalar en $\mathbb{R}^2$. Para $\alpha = 1$ y $\beta = 2$, hallar la matriz métrica G de este producto escalar en la base $(\bar{u}, \bar{v})$ con $\bar{u} = (1, 1)$ y $\bar{v} = (1, -1)$.

III.26. Sea V el espacio vectorial de las funciones reales definidas y continuas en el intervalo $[0, 1]$; en V se considera el producto escalar definido mediante

$$(f|g) = \int_0^1 f(x)g(x)\,dx$$

1. Hallar el valor de $(f|g)$ si f y g son las funciones dadas por:

$$f(x) = a_0 + a_1x + \cdots + a_nx^n$$

y

$$g(x) = b_0 + b_1x + \cdots + b_mx^m$$

2. Hallar un subespacio de dimensión $n \in \mathbb{N}$ de V que sea ortogonal a la función $f_0 \in V$ definida por $f_0(x) = x^n$.

III.27. Sea V un espacio vectorial real de dimensión $n \in \mathbb{N}$ y sea $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ una base cualquiera de V. Hallar, en esta base, la matriz métrica G del producto escalar con el que la nueva base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es ortonormal, siendo

$$\bar{e}_1 = \bar{u}_1,\ \bar{e}_2 = \bar{u}_1 + \bar{u}_2,\ \bar{e}_3 = \bar{u}_1 + \bar{u}_2 + \bar{u}_3,\ ...,\ \bar{e}_n = \\ = \bar{u}_1 + \bar{u}_2 + \cdots + \bar{u}_n$$

III.28. En un espacio vectorial V y respecto de una cierta base suya $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$, se define un producto escalar mediante

$$\bar{x} \cdot \bar{y} = x_1y_1 + \alpha x_2y_2 + 2x_3y_3 + x_1y_2 + x_2y_1 - \\ - x_1y_3 - x_3y_1 - x_2y_3 - x_3y_2$$

$$\bar{x} = x_1\bar{u}_1 + x_2\bar{u}_2 + x_3\bar{u}_3 \quad , \quad \bar{y} = y_1\bar{u}_1 + y_2\bar{u}_2 + y_3\bar{u}_3$$

donde $\alpha \in \mathbb{R}$ es el menor número natural posible. Se pide:

1. Hallar α.
2. Hallar un vector $\bar{a} \in V$ que forme ángulos iguales con los vectores de la base dada.
3. Hallar una base ortonormal $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ con la misma orientación que la base dada y tal que:
   - $\bar{e}_1$ tenga la dirección de $\bar{a}$ y sus coordenadas sean positivas.
   - $\bar{e}_2$ tenga iguales sus segunda y tercera coordenadas.
   - $\bar{e}_3$ tenga su primera coordenada positiva.

III.29. Sea V el espacio vectorial de los polinomios reales con una indeterminada, x, de grado menor o igual que dos, en el que se considera la base usual $(1, x, x^2)$. Considérese la aplicación $f: V \times V \to \mathbb{R}$,

$$f[p(x), q(x)] = \\ = \int_{-1}^{1} [ap(x)q(x) + bp'(x)q(x) + c(q'(x)p(x)]\,dx$$

donde $p'(x)$ denota al polinomio derivado de $p(x)$ y $a, b, c \in \mathbb{R}$ son fijos. Se pide:

1. Comprobar que f es una forma bilineal, determinando su matriz G en la base usual.
2. Hallar las relaciones entre a, b y c para que $p(x) * q(x) = f[p(x), q(x)]$ sea un producto escalar.
3. Para $a = 2$ y $b = c = 1/3$, hallar el subespacio U ortogonal al polinomio x, hallar una base ortogonal de U y hallar la proyección ortogonal sobre U del polinomio x^2.

III.30. En un espacio vectorial real V de dimensión 3, se considera una cierta base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$. Hallar la matriz métrica G de un producto escalar definido en V del que se sabe que:

- $\|\bar{e}_1\| = \sqrt{2}$ y $\|\bar{e}_2\| = \sqrt{3}$.
- $U = \{x_1\bar{e}_1 + x_2\bar{e}_2 + x_3\bar{e}_3 \in V / x_1 + x_2 + x_3 = 0\}$ es ortogonal a $\mathcal{V}(\bar{e}_1)$.
- La proyección ortogonal de $\bar{e}_1 + \bar{e}_2 + \bar{e}_3$ sobre $\bar{e}_2$ es $3\bar{e}_2$.

(Póngase la solución en función de cuantos parámetros se precise.)

III.31. En el espacio vectorial euclídeo canónico $\mathbb{R}^3$, sea $f: \mathbb{R}^3 \to \mathbb{R}^3$ la transformación resultante de aplicar, sucesivamente, la rotación de ángulo α alrededor de $\mathcal{V}(0, 0, 1)$, la simetría ortogonal respecto de $\mathcal{V}((1, 0, 0), (0, 1, 1))$ y la homotecia de razón $k \in \mathbb{R}$. Se pide:

1. Matriz A de f en la base canónica.
2. Los vectores que son proporcionales a sus transformados.

III.32. Sea V un espacio vectorial euclídeo tridimensional y sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ una base de la que se sabe que:

$$\bar{e}_1 \cdot \bar{e}_1 = 1,\ \bar{e}_2 \cdot \bar{e}_2 = 1,\ \bar{e}_3 \cdot \bar{e}_3 = 2,\ \bar{e}_1 \cdot \bar{e}_2 = 0$$

$\bar{a} = 2\bar{e}_2 - \bar{e}_3$ es ortogonal a $\bar{e}_1$ y a $\bar{e}_3$.

1. Hallar la matriz métrica G del producto escalar en la base dada.
2. Calcular una base ortonormal $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ de V.
3. Si $f: V \to V$ es la transformación definida por

$$3f(\bar{u}_1) = 2\bar{u}_1 - 2\bar{u}_2 + \bar{u}_3$$
$$3f(\bar{u}_2) = a\bar{u}_1 + \bar{u}_2 - 2\bar{u}_3$$
$$3f(\bar{u}_3) = b\bar{u}_1 + c\bar{u}_2 + 2\bar{u}_3$$

calcular los números reales a, b y c sabiendo que f es ortogonal. Hallar el ángulo que forman los vectores $f(\bar{e}_1)$ y $f(\bar{e}_2)$ y la matriz A' en f en la base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$.

III.33. Sea V un espacio vectorial euclídeo y sean U_1 y U_2 subespacios suplementarios de V. En U_1 hay definido un producto escalar $(*)$; en U_2 hay definido otro producto escalar $(\circ)$. Pruébese que existe un único producto escalar en V cuyas restricciones a U_1 y U_2 son los anteriores productos ($*$ y $\circ$) y que convierte a U_1 y U_2 en subespacios ortogonales.

III.34. En el espacio vectorial euclídeo canónico $\mathbb{R}^5$, se considera el subespacio U que forman los $(x_1, x_2, x_3, x_4, x_5)$ que verifican a:

$$\begin{cases} x_1 - x_2 + x_3 - x_5 = 0 \\ x_1 + x_2 - x_3 - 4x_4 - x_5 = 0 \\ 2x_1 - 3x_2 + x_3 - 2x_4 = 0 \end{cases}$$

Hallar una base ortonormal de U y otra del subespacio $U^\perp$, suplementario ortogonal de U. Hallar también las proyecciones ortogonales de $\bar{a} = (1, 0, 0, 0, 0)$ sobre U y sobre $U^\perp$.

III.35. En el espacio vectorial euclídeo canónico $\mathbb{R}^4$, se considera el subespacio

$$U = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 / x_1 + x_4 = x_2 + x_3\}$$

1. Hallar una base ortogonal de U.
2. De entre todos los vectores unitarios $\bar{a}$ que forman ángulo de 60° con (1, 0, 0, 0) y con (0, 1, 0, 0), hallar aquellos cuya proyección sobre U tiene la menor norma posible.

III.36. Sea V un espacio vectorial euclídeo, en el que se consideran dos subespacios cualesquiera U_1 y U_2. Se pide hallar las relaciones de inclusión que existen entre los siguientes subespacios:

1. $(U_1 + U_2)^\perp$ y $U_1^\perp \cap U_2^\perp$.
2. $(U_1 \cap U_2)^\perp$ y $U_1^\perp + U_2^\perp$.
3. $(U_1 \cap U_2)^\perp$ y $U_1^\perp + U_2^\perp$ si V tiene dimensión finita.

III.37. Sea V el espacio vectorial de las funciones reales definidas y continuas en el intervalo $[0, 2\pi]$, en el que se considera el producto escalar

$$(f|g) = \int_{-\pi}^{\pi} f(x)g(x)\,dx$$

Considérense las siguientes funciones $f_p \in V$ (para $p = 0, 1, \ldots, 2n$):

$$f_0(x) = a_0,\ f_1(x) = a_1 \operatorname{sen} x,\ f_2(x) = a_2 \cos x$$
$$f_3(x) = a_3 \operatorname{sen} 2x,\ f_4(x) = a_4 \cos 2x,\ \ldots,$$
$$f_{2n-1}(x) = a_{2n-1} \operatorname{sen} nx,\ f_{2n}(x) = a_{2n} \cos nx,\ (a_p \neq 0)$$

1. Comprobar que el sistema que forman estas funciones es ortogonal y hallar los coeficientes a_p para que sea ortonormal.
2. Hallar la proyección ortogonal sobre el subespacio que engendran dichas funciones de cada una de las dos funciones:

$$\varphi(x) = x \quad \text{y} \quad \psi(x) = x^2$$

III.38. Sea $f: V \to W$ una aplicación entre los espacios vectoriales euclídeos V y W. Compruébese que si f conserva el producto escalar, entonces f es lineal.

III.39. Sea $f: V \to W$ una aplicación entre los espacios vectoriales euclídeos V y W. Compruébese que si f es lineal y conserva las normas, entonces f conserva el producto escalar.

III.40. En el espacio vectorial V de las funciones continuas de $[-1, 1]$ en $\mathbb{R}$, considérense los dos productos escalares:

1. $(f|g) = \int_{-1}^{1} f(x)g(x)\,dx.$
2. $\langle f, g\rangle = \int_{-1}^{1} x^2 f(x)g(x)\,dx.$

Sea $F: V \to V$ la aplicación lineal definida por $F(f)(x) = xf(x)$ para $x \in [0, 1]$. Se pide:

1. Comprobar que $\langle,\rangle$ es efectivamente un producto escalar.
2. Analizar si F es una aplicación ortogonal del espacio euclídeo $(V, \langle, \rangle)$ en el $(V, (|))$.
3. Analizar si f es una aplicación biyectiva.

III.41. Sea V el espacio vectorial euclídeo canónico $\mathbb{R}^3$; sea $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ la base canónica de $\mathbb{R}^3$. Sea W el espacio vectorial de las matrices reales de tamaño 3×3, en el que se considera el producto escalar:

$$\langle A, B\rangle = \tfrac{1}{2} \operatorname{traza}(AB^t)$$

1. Analizar si existe alguna aplicación ortogonal

$f: V \to W$ tal que $f(\bar{e}_1) = M_1$, $f(\bar{e}_2) = M_2$ y $f(\bar{e}_3) = M_3$, siendo

$$M_1 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

$$M_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

2. Obtener la expresión de $f(x, y, z)$.

III.42. En el espacio vectorial euclídeo canónico $\mathbb{R}^n$, se considera un subespacio U; sea $U^\perp$ el suplementario ortogonal de U.

1. Si $f: \mathbb{R}^n \to \mathbb{R}^n$ es una transformación ortogonal, analizar si $f(U^\perp)$ es el suplementario ortogonal de $f(U)$.
2. Si $f_1: U \to U$ y $f_2: U^\perp \to U^\perp$ son dos transformaciones ortogonales, comprobar que existe una y sólo una transformación ortogonal $f: \mathbb{R}^n \to \mathbb{R}^n$ tal que sus restricciones a U y a $U^\perp$ son f_1 y f_2, respectivamente.

III.43. Sea V_2 un espacio vectorial euclídeo de dimensión 2 y sean $\bar{a}_1, \bar{a}_2, \bar{a}_3 \in V_2$ tales que cada dos son independientes y $\bar{a}_1 + \bar{a}_2 + \bar{a}_3 = \bar{o}$. Se considera un endomorfismo $f: V_2 \to V_2$ tal que $\|f(\bar{a}_i)\| = \|\bar{a}_i\|$ para $i = 1, 2, 3$. Pruébese que:

1. $f(\bar{a}_i) \cdot f(\bar{a}_j) = \bar{a}_i \cdot \bar{a}_j$ para $i = 1, 2, 3$ y $j = 1, 2, 3$.
2. f es una transformación ortogonal.

III.44. Sea $f: \mathbb{R}^3 \to \mathbb{R}^3$ el giro de 60° alrededor de la recta generada por el vector $\bar{u} = (1, 1, 1)$. Hallar:

1. $f(\bar{e}_1)$, $f(\bar{e}_2)$ y $f(\bar{e}_3)$, siendo $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ la base canónica de $\mathbb{R}^3$.
2. La matriz A de f en la base canónica, comprobando que es ortogonal directa.
3. La matriz B de f en cualquier base ortonormal $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ en la que $\bar{u}_3$ tenga la dirección y el sentido de $\bar{u} = (1, 1, 1)$.
4. El ángulo θ que debe formar un vector $\bar{v}$ con el vector $\bar{u} = (1, 1, 1)$ para que $\bar{v}$ y $f(\bar{v})$ formen un ángulo α dado, ¿cómo debe de ser α para que exista solución?

III.45. En el espacio vectorial euclídeo canónico $\mathbb{R}^3$, se considera el endomorfismo $f: \mathbb{R}^3 \to \mathbb{R}^3$ que, respecto de la base canónica, tiene por matriz a la

$$A = \begin{bmatrix} 0 & -1/\sqrt{3} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{3} & -1/\sqrt{6} \end{bmatrix}$$

1. Comprobar que f es una transformación ortogonal, analizando si es directa o inversa.
2. Hallar los vectores $\bar{u} \in \mathbb{R}^2$ tales que $f(\bar{u}) = \pm\bar{u}$.
3. Comprobar que f es producto de una rotación por una simetría respecto del plano ortogonal al eje de la rotación; hallar la simetría y la rotación.

III.46. Sabiendo que la siguiente matriz es ortogonal y que a y x son positivos, completar dicha matriz

$$\begin{bmatrix} 1/2 & 1/6 & a & x \\ 1/2 & 1/6 & b & y \\ 1/2 & 1/2 & 0 & z \\ 1/2 & -5/6 & 0 & t \end{bmatrix}$$

III.47. Sea $f: V_3 \to V_3$ una transformación ortogonal, en el espacio vectorial euclídeo V_3, de dimensión 3. Si la ecuación $f(\bar{u}) = \bar{u}$ no se verifica para ninguna $\bar{u} \in V_2$, pruébese que entonces f no es una rotación, pero sí lo es $f^2 = f \circ f$. Describir f^2 en función de f. Analizar si, para algún $n \in \mathbb{N}$, la transformación $f^n = f \circ f \circ \overset{n}{\cdots} \circ f$ puede ser la identidad.

III.48. Si $\bar{u}$, $\bar{v}$ y $\bar{w}$ son vectores de un espacio vectorial euclídeo de dimensión 3 orientado, demostrar que se verifica la siguiente relación (doble producto vectorial):

$$\bar{u} \wedge (\bar{v} \wedge \bar{w}) = (\bar{u} \cdot \bar{w})\bar{v} - (\bar{v} \cdot \bar{w})\bar{u}$$

III.49. Si $\bar{u}$, $\bar{v}$ y $\bar{w}$ son tres vectores de un espacio vectorial euclídeo de dimensión 3 orientado, demostrar que se verifica la siguiente relación (identidad de Jacobi):

$$\bar{u} \wedge (\bar{v} \wedge \bar{w}) + \bar{v} \wedge (\bar{w} \wedge \bar{u}) + \bar{w} \wedge (\bar{u} \wedge \bar{v}) = \bar{o}$$

III.50. Si $\bar{u}$, $\bar{v}$ y $\bar{w}$ son tres vectores de un espacio vectorial euclídeo de dimensión 3 orientado, comprobar que para el producto mixto $[\bar{u}, \bar{v}, \bar{w}]$ se verifica que:

$$[\vec{u}, \vec{v}, \vec{w}] = \begin{vmatrix} \|\vec{u}\|^2 & \vec{v} \cdot \vec{u} & \vec{w} \cdot \vec{u} \\ \vec{u} \cdot \vec{v} & \|\vec{v}\|^2 & \vec{w} \cdot \vec{v} \\ \vec{u} \cdot \vec{w} & \vec{v} \cdot w & \|\vec{w}\|^2 \end{vmatrix}$$

III.51. Sean $\vec{u}$ y $\vec{v}$ dos vectores del espacio vectorial euclídeo de dimensión 3 orientado V_3. Sea $f: V_3 \to V_3$ una transformación ortogonal. Hallar $f(\vec{u}) \wedge f(\vec{v})$ en función de $\vec{u} \wedge \vec{v}$.

III.52. Sean $\vec{u}$ y $\vec{v}$ dos vectores de un espacio vectorial euclídeo de dimensión 3 orientado. Demostrar que:

$$\|\vec{u} \wedge \vec{v}\|^2 + (\vec{u} \cdot \vec{v})^2 = \|\vec{u}\|^2 \|\vec{v}\|^2$$

III.53. En un espacio vectorial euclídeo V tridimensional, se considera la rotación $f: V \to V$ alrededor de la recta R engendrada por un cierto vector unitario $\vec{e}$ y de ángulo θ. Se pide:

1. Hallar la imagen $f(\vec{u})$ de un vector $\vec{u}$ ortogonal a $\vec{e}$.
2. Hallar la imagen $f(\vec{u})$ de un vector cualquiera $\vec{x} \in V$.
3. Si V es el espacio canónico $\mathbb{R}^3$, si $\theta = 45°$ y si $R = \mathbb{V}(2, 1, -2)$, hallar $f(\vec{x})$, donde $\vec{x}$ es un vector cualquiera de $\mathbb{R}^3$.

SOLUCIONES

III.1. Recúrrase a que $f(\bar{x}+\bar{y}, \bar{x}+\bar{y}) = 0$ para cualesquiera $\bar{x}, \bar{y} \in V$.

III.2. 1. Las primeras fila y columna de A quedan multiplicadas por λ (a_{11} pasa a $\lambda^2 a_{11}$).
2. Se permutan entre sí las primera y segunda fila y, también, las primera y segunda columnas.
3. A la primera fila se le suma λ veces la segunda y a la primera columna se le suma λ veces la segunda.

III.3. $C_1 \supset D_1$, $C_2 \supset D_2$, $C_1 = C_2 = V$.

III.4. Tómese una base en V; sea A la matriz de ω; sean M, M^*, N y N^* las matrices de φ, φ^*, ψ y ψ^*. Es:

$$M^* = A^{-1}M^tA \quad , \quad (MN)^* = N^*M^*$$
$$(\lambda M + \mu N)^* = \lambda M^* + \mu N^* \quad (M^*)^* = M$$
$$(\varphi \circ \psi)^* = \psi^* \circ \varphi^* \quad (\lambda\varphi + \mu\psi)^* = \lambda\varphi^* + \mu\psi^*$$
$$(\varphi^*)^* = \varphi$$

III.5. Recúrrase a que $\omega(\bar{u} \pm \bar{v}) = \omega(\bar{u}) + \omega(\bar{v}) \pm 2f(\bar{u}, \bar{v})$.

III.6 1. $f(\varphi, \psi) =$

$$\tfrac{1}{2}\int_{-a}^{a} [\varphi(x+b)\psi(x-b) + \varphi(x-b)\psi(x+b)]\,dx$$

2. $$A = \begin{bmatrix} 2a & 0 & \frac{2}{3}a^3 + 2ab^2 \\ 0 & \frac{2}{3}a^3 - 2ab^2 & 0 \\ \frac{2}{3}a^3 + 2ab^2 & 0 & \frac{2}{5}a^5 + 2ab^4 \end{bmatrix}$$

3. $N(\omega) = \mathcal{V}(x)$
Diagonal $(2\sqrt{3},\ 0,\ -12\sqrt{3}/5)$
Base de diagonalización $(1,\ x,\ -2 + x^2)$

III.7. M^tAM; para que $\omega \circ \varphi$ sea regular es necesario y suficiente que φ sea biyectiva (automorfismo) y ω sea regular.

III.8. Existe $\bar{v}$ no conjugado de $\bar{u}$. Para cualquier $\bar{x} \in V$, póngase $\bar{x} = \lambda\bar{v} + (\bar{x} - \lambda\bar{v})$ y hállese $\lambda \in K$ para que $\bar{x} - \lambda\bar{v}$ sea conjugado de $\bar{x}$.

III.9. D^c es el subespacio intersección de los subespacios conjugados de los vectores de D; esto es:

$$D^c = \bigcap_{\bar{x} \in D} \mathcal{C}(\bar{x})$$

III.10. $\omega(\bar{x}) = X^tAX = (X^tC)(FX) = X^t(CF)X$ y CF tiene rango 1.
Si rang $A = 1$, entonces $A = CF$ para ciertos $C \in \mathcal{M}_{n\times 1}$ y $F \in \mathcal{M}_{1\times n}$.

III.11. 2. $\omega(M) = 2x_2x_3 - 2x_1x_4$

3. $$A = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$$

4. Subespacio de las matrices simétricas.

III.12. 1. $f(p(x), q(x)) = \frac{1}{2}[p(\alpha)q(\beta) + p(\beta)q(\alpha)]$

2. $$A = \tfrac{1}{2}\begin{bmatrix} 2 & \alpha+\beta & \alpha^2+\beta^2 \\ \alpha+\beta & 2\alpha\beta & \alpha\beta(\alpha+\beta) \\ \alpha^2+\beta^2 & \alpha\beta(\alpha+\beta) & 2\alpha^2\beta^2 \end{bmatrix}$$

3. $\alpha + \beta = 0$
4. Si $\alpha \neq 0$, $N(\omega) = \mathcal{V}(\alpha^2 + x^2)$.
Si $\alpha = 0$, $N(\omega) = \mathcal{V}(x, x^2)$.
Diagonal $(1, \alpha^2, 0)$, base $(1, x, -\alpha^2 + x^2)$.

III.13. Sea $(\bar{u}_1, \ldots, \bar{u}_n)$ una de las bases en las que la matriz $A = [a_{ij}]$ de ω es diagonal. La base pedida puede ser la $(\bar{e}_1, \ldots, \bar{e}_n)$ siguiente:

$$\bar{e}_i = (1/\sqrt{a_{ii}})\bar{u}_i, \quad \text{si} \quad i \leq r$$
$$\bar{e}_i = \bar{u}_i \quad \text{si} \quad i > r$$

(nótese que, en general, $\sqrt{a_{ii}}$ será un número complejo).

III.14. 1. $$A = \begin{bmatrix} 1 & -2 & 1 \\ -2 & -2 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

2. $\omega(\bar{x}) = x_1^2 - x_2^2$ en la base $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$:

$\bar{u}_1 = (1, 0, 0)$

$\bar{u}_2 = (2/\sqrt{6}, 1/\sqrt{6}, 0)$

$\bar{u}_3 = (-1, 0, 1)$

III.15. 1. Diagonal $(1, \alpha - 1, \alpha + 1)$; base $((1, -1, 0), (2, -1, 0), (1, -1, 1))$

2. Si $\alpha = 1$, $N(\omega) = \mathcal{V}(2, -1, 0)$. Si $\alpha = -1$, $N(\omega) = \mathcal{V}(1, -1, 1)$. Si $\alpha \neq 1, -1$, entonces $N(\omega) = O$.

3. $(x - z) \pm \sqrt{2}(y + z) = 0$.

4. Base $((1, -1, 1), (\sqrt{2}, 1, 0), (\sqrt{2}, -1, 0))$, $\omega(\bar{x}) = -8x_2x_3$.

III.16. 1. $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & \alpha \end{bmatrix}$

2. Si $\alpha \neq 0$, $N(\omega) = O$.
Si $\alpha = 0$, $N(\omega) = \mathcal{V}(1, -1, 0, 1)$.

3. Diagonal $(1, 1, -1, \alpha)$.
Base $(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$, donde:

$\bar{u}_1 = (1, 0, 0, 0)$, $\bar{u}_2 = (0, 0, 1, 0)$

$\bar{u}_3 (-1, 1, -1, 0)$ $\bar{u}_4 = (1, -1, 0, 1)$

4. $(\bar{u}_1 + \bar{u}_3, \bar{u}_2 + \bar{u}_3, 2\bar{u}_3 + \bar{u}_4, \bar{u}_1 + 2\bar{u}_2 + 3u_3 + u_4)$

$\omega(\bar{x}) = -2x_1x_2 - 4x_1x_3 - 4x_1x_4 - 4x_2x_3 - 2x_2x_4 - 4x_3x_4$

III.17. Supóngase $\omega(\bar{u}) > 0$ y $\omega(\bar{v}) < 0$. Búsquese un $\bar{x} = \lambda\bar{u} + \bar{v}$, $\bar{x} \neq \bar{o}$, tal que $\omega(\bar{x}) = 0$.

III.18. $(M^tM)^t = M^tM$. Si $X \neq O$ es matriz columna, $MX \neq O$ y, por ello, $(MX)^t(MX) > 0$ si $X \neq O$; luego $X^t(M^tM)X > 0$ para todo $X \neq O$.

III.19. La matriz A ha de ser simétrica y definida positiva.

III.20. $x_1 = y_1 - \dfrac{1}{a_{11}}(a_{12}y_2 + \cdots + a_{1n}y_n)$; $x_i = y_i$ para $i = 2, 3, \ldots, n$.

$x_1 = y_1 + y_2$, $x_2 = y_1 - y_2$, $x_i = y_i$ para $i = 3, \ldots, n$.

III.21. Recúrrase a que $\omega(\lambda\bar{u} + \bar{v}) \geq 0$, $\forall \lambda \in \mathbb{R}$.

III.22. 1. $A = \begin{bmatrix} c-1 & 0 & 0 & -1 \\ 0 & 0 & c & 0 \\ 0 & c & 0 & 0 \\ -1 & 0 & 0 & c-1 \end{bmatrix}$

2. Diagonalización para $c = 1$:
Diagonal: $-2, 2, 2, -2$.
Base:

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Diagonalización para $c \neq 1$:
Diagonal: $c - 1, 2c, -2c, c(c-1)(c-2)$.
Base:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1-c \end{bmatrix}$$

3. Nunca es definida; semidefinida negativa si $c = 0$.

III.23. 1. Si $\alpha = 0$, diagonal $(2, 6, 2, -2)$ y base $(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$ donde:

$\bar{u}_1 = (1, 0, 0, 0)$, $\bar{u}_2 = (-1, 0, 2, 0)$

$\bar{u}_3 = (0, 1, 0, 1)$, $\bar{u}_4 = (0, -1, 0, 1)$

Si $\alpha \neq 0$, diagonal $(2, 6, \alpha, (\alpha^2 - 1)\alpha)$ y base $(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4)$ donde:

$\bar{u}_1 = (1, 0, 0, 0)$, $\bar{u}_2 = (-1, 0, 2, 0)$

$\bar{u}_3 = (0, 1, 0, 0)$, $\bar{u}_4 = (0, -1, 0, \alpha)$

2. Si $|\alpha| < 1$, rang $\omega = 4$ y sig $\omega = (3, 1)$.
Si $\alpha > 1$, rang $\omega = 4$ y sig $\omega = (4, 0)$.
Si $\alpha = 1$, rang $\omega = 3$ y sig $\omega = (3, 0)$.
Si $\alpha = -1$, rang $\omega = 3$ y sig $\omega = (2, 1)$.
Si $\alpha < -1$, rang $\omega = 4$ y sig $\omega = (2, 2)$.

3. ω definida positiva sii $\alpha > 1$; ω semidefinida positiva sii $\alpha = 1$.

III.24. Según el criterio de Sylvester ha de ser $\alpha > 0$. Diagonalizando por congruencia se obtiene (entre otras muchas) la base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$:

$$[G|I] = \left[\begin{array}{ccc|ccc} 5 & -2 & 2 & 1 & 0 & 0 \\ -2 & 2 & -1 & 0 & 1 & 0 \\ 2 & -1 & 2 & 0 & 0 & 1 \end{array}\right] \simeq$$

$$\simeq \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 & 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1 & 3/\sqrt{21} & 2/\sqrt{21} & -2/\sqrt{21} \end{array}\right]$$

$$\bar{e}_1 = \left(0, 0, \frac{1}{\sqrt{2}}\right) \quad ; \quad \bar{e}_2 = \left(0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

$$\bar{e}_3 = \left(\frac{3}{\sqrt{21}}, \frac{2}{\sqrt{21}}, \frac{-2}{\sqrt{21}}\right)$$

III.25. Según el criterio de Sylvester ha de ser $\alpha \neq \beta$.

$$G = \begin{bmatrix} 26 & -14 \\ -14 & 10 \end{bmatrix}$$

III.26. 1. $(f|g) = \sum_{i=0}^{n} \sum_{j=0}^{m} \frac{a_i b_j}{1+i+j}$

2. Las siguientes n funciones $f_p \in V$ son ortogonales a f_0; el subespacio que engendran es solución:

$$f_p(x) = (n+p)x^{p-1} - (n+p+1)x^p$$

$p = 1, 2, \ldots, n$.

III.27.

$$G = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 2 \end{bmatrix}$$

III.28. 1. Según el criterio de Sylvester, $\alpha = 2$.

2. $\bar{a} = 3\bar{u}_1 + (\sqrt{2}-1)\bar{u}_2 + (\sqrt{2}+1)\bar{u}_3$.

3. $$\bar{e}_1 = \frac{1}{\sqrt{70}}(6\bar{u}_1 + 3\bar{u}_2 + 5\bar{u}_3)$$

$$\bar{e}_2 = \frac{1}{\sqrt{6}}(-2\bar{u}_1 + \bar{u}_2 + \bar{u}_3)$$

$$\bar{e}_3 = \frac{1}{\sqrt{195}}(11\bar{u}_1 - 7\bar{u}_2 + 5\bar{u}_3)$$

III.29. 1. $$G = \frac{1}{15}\begin{bmatrix} 30a & 30c & 10a \\ 30b & 10a & 10b+20c \\ 10a & 20b+10c & 6a \end{bmatrix}$$

2. $b = c$, $a > 0$, $a^2 > 18b^2$.

3. $U = \{\alpha + \beta x + \gamma x^2 / \alpha + 2\beta + \gamma = 0\}$

una base ortogonal de U:

$(1 - x^2, 1 + 2x - 5x^2)$

Proyección: $-\frac{1}{2}x + x^2$.

III.30. $$G = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 3 & 4 \\ 2 & 4 & \alpha \end{bmatrix} \quad \text{para} \quad \alpha > 6$$

III.31. 1. $$A = \begin{bmatrix} k\cos\alpha & -k\,\text{sen}\,\alpha & 0 \\ 0 & 0 & k \\ k\,\text{sen}\,\alpha & k\cos\alpha & 0 \end{bmatrix}$$

2. $(\lambda\cos\alpha, \lambda(\cos\alpha + 1), -\lambda(\cos\alpha + 1))$, $\lambda \in \mathbb{R}$

$(0, \mu, -\mu)$, $\mu \in \mathbb{R}$

(a, b, b) , $a, b \in \mathbb{R}$

III.32. 1. $$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

2. $\bar{u}_1 = \bar{e}_1$, $\bar{u}_2 = \bar{e}_2$, $\bar{u}_3 = \bar{e}_3 - \bar{e}_2$

3. $a = 2$, $b = 1$, $c = 2$

$$\text{ang}\,(f(\bar{e}_1), f(\bar{e}_2)) = 90°$$

$$A' = \frac{1}{3}\begin{bmatrix} 2 & 2 & 3 \\ -3 & 3 & 3 \\ 1 & -2 & 0 \end{bmatrix}$$

III.33. Dados $\bar{x}, \bar{y} \in V$, se puede poner de manera única $\bar{x} = \bar{x}_1 + \bar{x}_2$ e $\bar{y} = \bar{y}_1 + \bar{y}_2$ con $\bar{x}_1, \bar{y}_1 \in U_1$ y $\bar{x}_2, \bar{y}_2 \in U_2$. El producto escalar pedido $(\cdot)$ está definido por $\bar{x} \circ \bar{y} = \bar{x}_1 * \bar{y}_1 + \bar{x}_2 \circ \bar{y}_2$.

III.34. Una de las muchas bases ortonormales de U y de $U^\perp$ son las $(\bar{u}_1, \bar{u}_2)$ y $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ formadas por:

$$\bar{u}_1 = (1/2, 1/2, 1/2, 0, 1/2)$$

$$\bar{u}_2 = (2/3, 0, -2/3, 1/3, 0)$$

$$\bar{e}_1 = (0, 1/\sqrt{2}, 0, 0, -1/\sqrt{2})$$

$$\bar{e}_2 = (-1/3\sqrt{2}, 0, 1/3\sqrt{2}, 4/3\sqrt{2}, 0)$$

$$\bar{e}_3 = (-1/2, 1/2, -1/2, 0, 1/2)$$

Proyección sobre U:

$$\bar{a}' = \frac{1}{2}\bar{u}_1 + \frac{2}{3}\bar{u}_2 =$$

$$= \left(\frac{25}{36}, \frac{9}{36}, \frac{-7}{36}, \frac{8}{36}, \frac{9}{36}\right)$$

Proyección sobre $U^\perp$:

$$\bar{a} - \bar{a}' = \left(\frac{11}{36}, \frac{-9}{36}, \frac{7}{36}, \frac{-8}{36}, \frac{-9}{36}\right)$$

III.35. 1. Una de las muchas soluciones puede ser la $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ con

$$\bar{e}_1 = (1, 0, 1, 0)$$

$$\bar{e}_2 = (0, 1, 0, 1)$$

$$\bar{e}_3 = (1, 1, -1, -1)$$

2. $\bar{a} = (1/2, 1/2, \cos\alpha/\sqrt{2}, \operatorname{sen}\alpha/\sqrt{2})$ ha de ser tal que haga mínimo a

$$\left\| \frac{1+\sqrt{2}\cos\alpha}{4}\bar{e}_1 + \frac{1+\sqrt{2}\operatorname{sen}\alpha}{4}\bar{e}_2 + \right.$$

$$\left. + \frac{2-\sqrt{2}\cos\alpha-\sqrt{2}\operatorname{sen}\alpha}{8}\bar{e}_3 \right\| =$$

$$= \frac{7+\operatorname{sen}2\alpha}{8}$$

lo que ocurre para $\alpha = 3\pi/4$ y $\alpha = -\pi/4$, luego

$$\bar{a} = (\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2})$$

o

$$\bar{a} = (\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2})$$

III.36. 1. $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$

2. $(U_1 \cap U_2)^\perp \supset U_1^\perp + U_2^\perp$

3. $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$

III.37. 1. $a_0 = 1/\sqrt{2\pi}$, $a_p = 1/\sqrt{\pi}$ $(p > 1)$

2. Proyección de $\psi(x)$:

$$\frac{2}{1}\operatorname{sen}x - \frac{2}{2}\operatorname{sen}2x + \frac{2}{3}\operatorname{sen}3x - \cdots +$$

$$+ (-1)^{n-1}\frac{2}{n}\operatorname{sen}nx$$

Proyección de $\psi(x)$:

$$\frac{\pi^2}{3} - \frac{\cos x}{1^2} + \frac{\cos 2x}{2^2} - \frac{\cos 3x}{3^2} + \cdots +$$

$$+ (-1)^n \frac{\cos nx}{n^2}$$

III.38. Compruébese que la norma del vector $f(\lambda\bar{x} + \mu\bar{y}) - \lambda f(\bar{x}) - \mu f(\bar{y})$ es igual a la norma de $(\lambda\bar{x} + \mu\bar{y}) - \lambda\bar{x} - \mu\bar{y}$ y que, por tanto, es nula.

III.39. Compruébese que $4(\bar{a} \cdot \bar{b}) = \|\bar{a} + \bar{b}\|^2 - \|\bar{a} - \bar{b}\|^2$; tómese $\bar{a} = \bar{x}$ y $\bar{b} = \bar{y}$; tómese después $\bar{a} = f(\bar{x})$ y $\bar{b} = f(\bar{y})$.

III.40. 2. F es ortogonal (lineal y conserva el producto escalar).

3. F no es sobreyectiva; $f(x) = 1$ no es imagen por F de ninguna función continua.

III.41. 1. Como $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ es una base ortonormal y (M_1, M_2, M_3) es un sistema ortonormal (se comprueba fácilmente), existe la aplicación ortogonal f.

2.

$$f(x, y, z) = \begin{bmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{bmatrix}$$

III.42. 1. Si, ya que $f(U) \perp f(U^\perp)$ y $\dim f(U) + \dim f(U^\perp) = \dim U + \dim U^\perp = n$.

2. $\forall \bar{x} \in \mathbb{R}^n$, $\bar{x} = \bar{x}_1 + \bar{x}_2$ con $\bar{x}_1 \in U$ y $\bar{x}_2 \in U^\perp$ (descomposición única); $f(\bar{x}) = f_1(\bar{x}_1) + f_2(\bar{x}_2)$.

III.43. 1. Aplicar el «teorema del coseno» de los triángulos ($a^2 = b^2 + c^2 - 2bc \cos A$).

2. Refiérase f a las bases $(\bar{a}_1, \bar{a}_2)$, en V_2 origen, y $(f(\bar{a}_1), f(\bar{a}_2))$ en V_2 imagen; la matriz de f es la identidad; las matrices métricas son iguales.

III.44. 1.
$$f(\bar{e}_1) = (2/3, 2/3, -1/3)$$
$$f(\bar{e}_2) = (-1/3, 2/3, 2/3)$$
$$f(\bar{e}_3) = (2/3, -1/3, 2/3)$$

2. $$A = \begin{bmatrix} 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \\ -1/3 & 2/3 & 2/3 \end{bmatrix} \quad A^t A = I, \ \det A = 1$$

3. $$B = \begin{bmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

4. $\cos \theta = \sqrt{2 \cos \alpha - 1}$; $\alpha \leq 60°$

III.45. 1. $A^t A = I$ y $\det A = -1$, luego f es ortogonal inversa.

2. $\bar{u} = \rho(\sqrt{3} - \sqrt{2}, \ 1 + \sqrt{2} - \sqrt{6}, \ 1 - \sqrt{2})$ para $\rho \in \mathbb{R}$.

3. El eje de giro es la recta engendrada por $\bar{u}$.
El plano de la simetría es el ortogonal a $\bar{u}$.
El ángulo de giro α es el ángulo áng $(\bar{v}, f(\bar{v}))$, donde $\bar{v}$ es ortogonal a $\bar{u}$, por ejemplo $\bar{v} = (\sqrt{2}, 1, 1)$

$$\cos \alpha = \frac{6 + 2\sqrt{3} - \sqrt{6}}{12} \simeq 0{,}585 \quad , \quad \alpha \simeq 54{,}23°$$

III.46. $a = 1/\sqrt{2}$, $b = -1/\sqrt{2}$, $x = 1/2$, $z = -3/4$, $t = -1/4$.

III.47. f es la composición de una rotación, de cierto ángulo α y alrededor de cierta recta R, seguido de la simetría respecto del plano perpendicular a R; f^2 es la rotación de ángulo 2α alrededor de R. La transformación f^n es una rotación de ángulo $n\alpha$; ha de ser $2\pi/\alpha$ racional.

III.48. Poner $\bar{u} = \lambda\bar{v} + \mu\bar{w} + \nu(\bar{v} \wedge \bar{w})$ y recurrir al último ejercicio del Capítulo 8.

III.49. Recúrrase a la fórmula del problema anterior.

III.50. Recurrir a una base ortonormal, poner $[\bar{u}, \bar{v}, \bar{w}]$ en forma del determinante de la matriz de coordenadas y recurrir a que $[\bar{u}, \bar{v}, \bar{w}]^2$ es el determinante de dicha matriz por su traspuesta.

III.51. $f(\bar{u}) \wedge f(\bar{v}) = (\det f)\,(\bar{u} \wedge \bar{v})$.

III.52. Recurrir a coordenadas en una base ortonormal directa.

III.53. 1. $f(\bar{u}) = \cos \theta \bar{u} - \operatorname{sen} \theta(\bar{e} \wedge \bar{u})$

2. $f(\bar{x}) = (\bar{x} \cdot \bar{e})\bar{e} - \cos \theta \, [\bar{x} - (\bar{x} \cdot \bar{e})\bar{e}] + \operatorname{sen} \theta(\bar{e} \wedge \bar{x})$

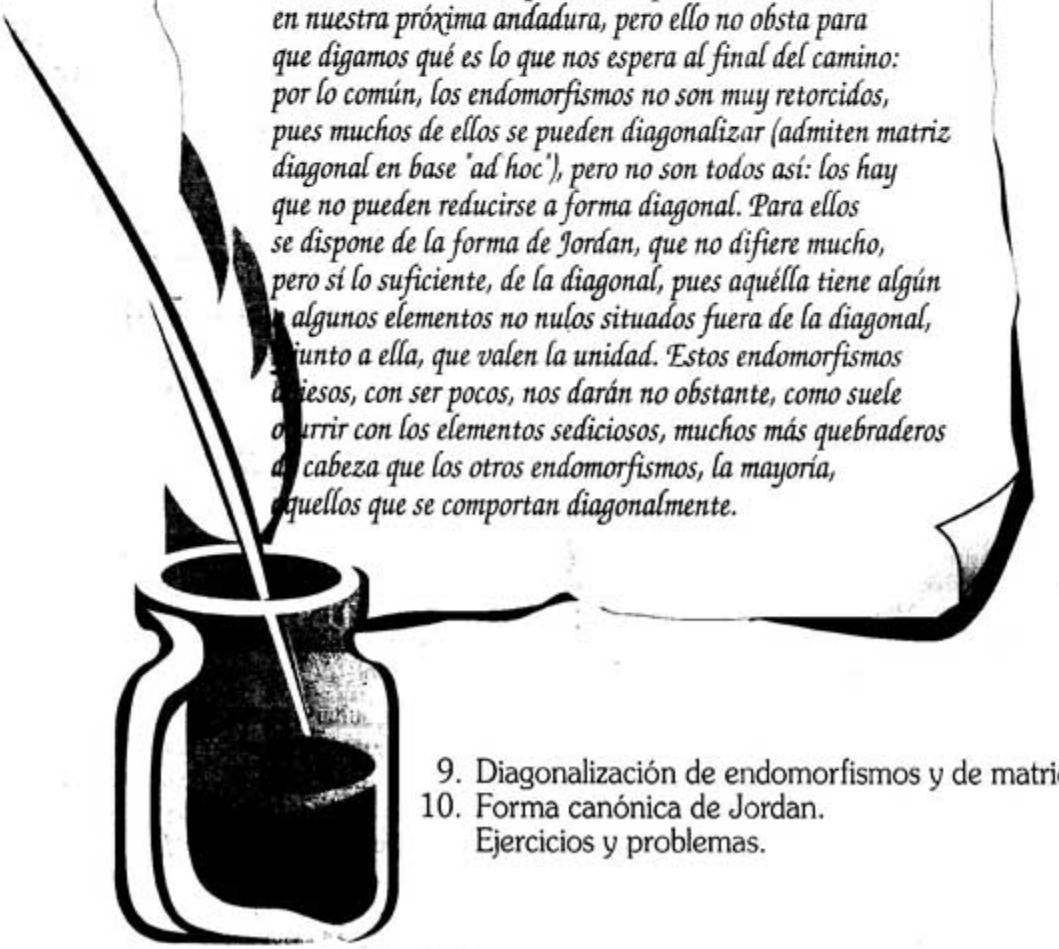
3. Si $\bar{x} = (x_1, x_2, x_3)$ y $f(\bar{x}) = (x'_1, x'_2, x'_3)$, es

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \frac{1}{18} \begin{bmatrix} 8 + 5\sqrt{2} & 4 + 4\sqrt{2} & -8 + 7\sqrt{2} \\ 4 - 8\sqrt{2} & 2 + 8\sqrt{2} & -4 - 4\sqrt{2} \\ -8 + \sqrt{2} & -4 + 8\sqrt{2} & 8 + 5\sqrt{2} \end{bmatrix}$$

PARTE IV

INTERESADOS, como estamos, y mucho, en conocer los asuntos de la linealidad, no podemos por menos de ahondar en los entresijos de los endomorfismos, con el propósito de desvelar su estructura, de contemplar sus entrañas, despojándoles de todo aquello que no les es propio o esencial, lo que nos llevará a referirles a bases íntimamente ligadas a ellos. Y, en este afán, nos daremos de bruces con los autovalores y los autovectores, también llamados valores (escalares) y vectores propios, que han de sernos de enorme utilidad en nuestra tarea.

No es cosa de contar aquí todo lo que nos va a ocurrir en nuestra próxima andadura, pero ello no obsta para que digamos qué es lo que nos espera al final del camino: por lo común, los endomorfismos no son muy retorcidos, pues muchos de ellos se pueden diagonalizar (admiten matriz diagonal en base "ad hoc"), pero no son todos así: los hay que no pueden reducirse a forma diagonal. Para ellos se dispone de la forma de Jordan, que no difiere mucho, pero sí lo suficiente, de la diagonal, pues aquélla tiene algún o algunos elementos no nulos situados fuera de la diagonal, junto a ella, que valen la unidad. Estos endomorfismos díscolos, con ser pocos, nos darán no obstante, como suele ocurrir con los elementos sediciosos, muchos más quebraderos de cabeza que los otros endomorfismos, la mayoría, aquellos que se comportan diagonalmente.



9. Diagonalización de endomorfismos y de matrices.
10. Forma canónica de Jordan. Ejercicios y problemas.

CAPITULO

9

Diagonalización de endomorfismos y de matrices

Cuando, hace poco, estudiábamos las aplicaciones lineales, de un espacio vectorial V en otro W (ambos de dimensión finita), vimos que, dada una tal aplicación $f: V \to W$, había ciertas bases en V y en W respecto de las cuales la matriz de f es diagonal. Estas bases no son únicas y se pueden hallar fácilmente, por ejemplo, mediante operaciones elementales. Ahora que vamos a considerar el caso $V = W$ y va a haber una sola base, la situación es más compleja y no tiene siempre una solución tan sencilla.

Se trata aquí de lo siguiente: dado un endomorfismo $f: V \to V$ (donde V es un espacio vectorial de dimensión finita) y llamando A a la matriz de f en cierta base de V, vamos a estudiar si al cambiar de base se puede conseguir una matriz A' de f (nótese que A' es semejante a A: esto es, $A' = P^{-1}AP$, donde P es la matriz de cambio de coordenadas; véase [094]) que sea sencilla, cuanto más mejor; o sea, de modo que A' tenga el mayor número posible de ceros. En muchos casos se puede conseguir que A' sea diagonal; se dice entonces que f es diagonalizable. Cuando no es así, hay que conformarse con otras formas reducidas menos simples; de esto último hablaremos en el próximo capítulo.

α AUTOVALORES Y AUTOVECTORES

Para poder estudiar la estructura interna de los endomorfismos y, en particular, su reducción a forma de Jordan o a forma diagonal, cuando ello sea posible, vamos a precisar de los «autovalores» y los «autovectores», a los que dedicamos la primera parte de este capítulo.

9.1. AUTOVALORES Y AUTOVECTORES. POLINOMIO CARACTERISTICO

Aun cuando a primera vista pueda resultar sorprendente, en el estudio de las cuestiones que aquí nos ocupan, referentes a un endomorfismo f, desempeñan un papel muy importante aquellos vectores $\bar{x}$, no nulos, tales que $f(\bar{x})$ es proporcional a $\bar{x}$. Se dice que tales $\bar{x}$ son autovectores de f y que la constante de proporcionalidad, entre $f(\bar{x})$ y $\bar{x}$, es un autovalor de f:

AUTOVALORES Y AUTOVECTORES DE UN ENDOMORFISMO

[161]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$; sea K el cuerpo de los escalares. Se dice que un escalar $\lambda \in K$ es un *autovalor* (o valor propio) de f si existe algún vector no nulo $\bar{x} \in V$ tal que $f(\bar{x}) = \lambda\bar{x}$, esto es, si el endomorfismo $f - \lambda i$ (donde i la aplicación idéntica de V en V) no es inyectivo.

Si λ es un autovalor de f, los vectores $\bar{x} \in V$ tales que $f(\bar{x}) = \lambda\bar{x}$ se llaman *autovectores* (o vectores propios) de f asociados a λ; estos autovectores forman un subespacio, V_λ, que es:

$$V_\lambda = \{\bar{x} \in V \,/\, f(\bar{x}) = \lambda\bar{x}\} \quad , \quad \text{esto es } V_\lambda = \text{Nuc}\,(f - \lambda i)$$

A V_λ se le llama *subespacio propio* del endomorfismo f asociado a su autovalor λ.

OBSERVACIONES

1. Si se verifica la anterior relación $f(\bar{x}) = \lambda\bar{x}$, para poder decir que el escalar $\lambda \in \mathbb{R}$ es un autovalor de f se ha exigido que sea $\bar{x} \neq \bar{o}$. Si se hubiera admitido la posibilidad $\bar{x} = \bar{o}$, habría resultado que todo escalar λ sería autovalor, ya que $f(\bar{o}) = \bar{o} = \lambda\bar{o}$ para cualquiera que sea $\lambda \in K$.
2. Dado un escalar $\lambda \in K$, es obvio que cada una de las siguientes relaciones equivale a la que le sigue:

$$\bar{x} \in V_\lambda, \quad f(\bar{x}) = \lambda\bar{x}, \quad f(\bar{x}) - \lambda\bar{x} = \bar{o}$$

$$f\,(\bar{x}) - \lambda i(x) = \bar{o}, \quad (f - \lambda i)(\bar{x}) = \bar{o}, \quad \bar{x} \in \text{Nuc}\,(f - \lambda i)$$

   De ahí que se verifiquen las anteriores afirmaciones: 1) $\lambda \in K$ es autovalor de f si y sólo si $f - \lambda i$ no es inyectiva; y 2) V_λ es el núcleo de $f - \lambda i$.
3. El escalar $\lambda = 0$ es autovalor de f si y sólo si $f - 0i$ no es inyectivo, es decir, si f no es inyectivo.
4. Un endomorfismo cualquiera no tiene que tener, necesariamente, algún autovalor. Así ocurre con los dos ejemplos siguientes:

   - Si V es el espacio vectorial de los polinomios, $p(x)$, reales o complejos, el endomorfismo $p(x) \to xp(x)$ carece de autovalores.
   - El endomorfismo $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (-y, x)$, tampoco tiene autovalores, ya que

$$\left.\begin{array}{l} f(x, y) = \lambda(x, y) \\ (x, y) \neq (0, 0) \end{array}\right\} \Rightarrow \left.\begin{array}{l} y = -\lambda x \ , \ x = \lambda y \\ (x, y) \neq (0, 0) \end{array}\right\} \Rightarrow \begin{array}{l} y = -\lambda^2 y \\ y \neq 0 \end{array} \Rightarrow \lambda^2 = -1$$

   y no existe ningún $\lambda \in \mathbb{R}$ tal que $\lambda^2 = -1$.

5. Al conjunto de los autovalores de un endomorfismo f se le llama *espectro* de f, y se le denota por $\sigma(f)$. Nótese que la relación $\lambda \notin \sigma(f)$ equivale a la $V_\lambda = O$.

EJERCICIO

Sea V el espacio vectorial de los polinomios complejos (o reales) con una indeterminada x; sea $f: V \to V$ el siguiente endomorfismo:

$$p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n \mapsto f(p(x)) = a_n + a_{n-1}x + \cdots + a_1x^{n-1} + a_0x^n$$

Hallar los autovalores y los autovectores de f.

RESOLUCION

La ecuación $f(p(x)) = \lambda p(x)$, donde $p(x) \in V$ y $\lambda \in \mathbb{C}$, equivale a las:

$$a_n = \lambda a_0, \ldots, a_i = \lambda a_{n-i}, \ldots, a_0 = \lambda a_n$$

De $a_i = \lambda a_{n-i}$ y de $a_{n-i} = \lambda a_i$, se deduce que $a_i = \lambda^2 a_i$ para $i = 0, 1, \ldots, n$. Como ha de ser $p(x) \neq o$, alguno de los a_i deben ser no nulos y, dado que $a_i = \lambda^2 a_i$, resulta que $\lambda^2 = 1$. Los autovalores de f son, pues, $\lambda = 1$ y $\lambda = -1$ y los correspondientes subespacios propios serán:

$$V_1 = \{p(x) \in V \,/\, a_i = a_{n-i} \quad \text{para } i = 0, 1, \ldots, n\}$$
$$V_{-1} = \{p(x) \in V \,/\, a_i = -a_{n-i} \quad \text{para } i = 0, 1, \ldots, n\}$$

PROPOSICION

[162]

Sean $\lambda_1, \lambda_2, \ldots, \lambda_p \in K$ autovalores distintos de un endomorfismo $f: V \to V$, del espacio vectorial $V \neq O$ (K es el cuerpo de escalares); se verifica que:

1. Si $\bar{x}_i \neq \bar{o}$ es un vector propio de f correspondiente al autovalor λ_i para $i = 1, 2, \ldots, p$, entonces $\{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p\}$ es un sistema linealmente independiente.
2. Los subespacios propios V_{λ_i} asociados a los autovalores λ_i, para $i = 1, 2, \ldots, p$, son independientes, es decir, su suma es directa:

$$V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_p} = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_p}$$

DEMOSTRACION

1. Supongamos que la proposición es falsa, es decir, que $\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_p\}$ es linealmente dependiente. Sea $\bar{x}_{r+1}$ el primero de los vectores del sistema que depende linealmente de los que le preceden; es decir, $\bar{x}_1, ..., \bar{x}_r$ son linealmente independientes y

$$\bar{x}_{r+1} = h_1\bar{x}_1 + h_2\bar{x}_2 + \cdots + h_r\bar{x}_r \quad \text{(para ciertos escalares } h_i\text{)} \qquad [1]$$

Hallando la imagen por f de ambos miembros de esta igualdad, como $f(\bar{x}_i) = \lambda_i\bar{x}_i$ para $i = 1, 2, ..., r+1$, resulta que:

$$\lambda_{r+1}\bar{x}_{r+1} = h_1\lambda_1\bar{x}_1 + h_2\lambda_2\bar{x}_2 + \cdots + h_r\lambda_r\bar{x}_r \qquad [2]$$

Multiplicando ambos miembros de [1] por λ_{r+1} y restándole la [2], resulta que

$$\bar{o} = h_1(\lambda_{r+1} - \lambda_1)\bar{x}_1 + h_2(\lambda_{r+1} - \lambda_2)\bar{x}_2 + \cdots + h_r(\lambda_{r+1} - \lambda_r)\bar{x}_r$$

Ahora bien, como $\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_r\}$ es linealmente independiente, de aquí se deduce que los coeficientes $h_i(\lambda_{r+1} - \lambda_i)$ han de ser todos nulos (para $i = 1, 2, ..., r$) y, dado que los autovalores son distintos ($\lambda_{r+1} \neq \lambda_i$ para $i = 1, 2, ..., r$), ha de ser $h_1 = h_2 = \cdots = h_r = 0$, luego de [1] se desprendería que $\bar{x}_{r+1} = \bar{o}$, que va contra la hipótesis. Esta contradicción obliga a rechazar la no independencia de $\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_p\}$.

2. Supongamos que la suma no fuese directa, es decir (véase [074], II), que existen $\bar{x}_1 \in V_{\lambda_1}, ..., \bar{x}_p \in V_{\lambda_p}$ no todos nulos y tales que $\bar{x}_1 + \cdots + \bar{x}_p = \bar{o}$. Podemos suponer, sin pérdida de generalidad, que los vectores no nulos son los r primeros ($1 \leq r \leq p$), con lo que $\bar{x}_1 + \cdots + \bar{x}_r = \bar{o}$, siendo $\bar{x}_i \in V_{\lambda_i}$ y $\bar{x}_i \neq \bar{o}$ para $i = 1, 2, ..., r$, y esto no es posible pues contradice al resultado precedente. Así, pues, la suma del enunciado tiene que ser directa.

OBSERVACIONES

1. Si λ_1 y λ_2 son autovalores distintos de un endomorfismo, entonces sus subespacios propios V_{λ_1} y V_{λ_2} tiene intersección nula, $V_{\lambda_1} \cap V_{\lambda_2} = O$, ya que su suma es directa.
2. Si $f: V \to V$ es un endomorfismo en un espacio vectorial V de dimensión finita n, entonces f no puede tener más de n autovalores distintos. En efecto, si tuviese más de n, según la anterior propiedad 1 de [162], en V habría un sistema de más de n vectores que sería linealmente independiente, lo que no es posible.

AUTOVALORES Y AUTOVECTORES DE UNA MATRIZ CUADRADA

[163]

Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K. Se dice que $\lambda \in K$ es un *autovalor* (o valor propio) de A si existe alguna matriz columna no nula $X \in \mathcal{M}_{n\times 1}$ tal que $AX = \lambda X$. Si λ es un autovalor de A, las columnas X tales que $AX = \lambda X$ se llaman *autovectores* (o vectores propios) de A asociados a λ; todos los vectores propios asociados a λ forman un subespacio vectorial de $\mathcal{M}_{n\times 1}$, que se llama *subespacio propio* de A asociado al autovalor λ.

Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$; sea K el cuerpo de escalares. Si $\dim V = n$ y si A es la matriz de f en una cierta base de V, entonces:

1. $$\begin{pmatrix}\text{El escalar } \lambda \in K\\ \text{es autovalor de } f\end{pmatrix} \Leftrightarrow \begin{pmatrix}\text{El escalar } \lambda \in K\\ \text{es autovalor de } A\end{pmatrix} \Leftrightarrow$$
$$\Leftrightarrow \begin{pmatrix}\text{El escalar } \lambda \in K \text{ es tal que}\\ A - \lambda I \text{ es una matriz singular}\end{pmatrix}$$

2. Si λ es un autovalor de f (o de A) y llamando X a la columna de coordenadas de un vector $\bar{x} \in V$, en la base citada, entonces:
$$\begin{pmatrix}\bar{x} \text{ es un autovector}\\ \text{de } f \text{ asociado a } \lambda\end{pmatrix} \Leftrightarrow \begin{pmatrix}X \text{ es un autovector}\\ \text{de } A \text{ asociado a } \lambda\end{pmatrix} \Leftrightarrow (A - \lambda I)X = O$$

La dimensión del subespacio propio V_λ (de f o de A) asociado a λ es:

$$\dim V_\lambda = n - \operatorname{rang}(A - \lambda I)$$

COMPROBACIONES

1. Decir que $\lambda \in K$ es autovalor de f significa que es $f(\bar{x}) = \lambda\bar{x}$ para algún $\bar{x} \in V$ no nulo. Utilizando coordenadas, esto último se expresa diciendo que se verifica $AX = \lambda X$ para alguna $X \in \mathcal{M}_{n\times 1}$ no nula, que a su vez equivale a decir que λ es un autovalor de A.

   Nótese ahora que la relación $AX = \lambda X$ puede ponerse en la forma $AX = \lambda IX$ (donde I es la matriz unidad), o en la $(A - \lambda I)X = O$. Por tanto, $\lambda \in K$ será autovalor de A si, y sólo si, la ecuación matricial $(A - \lambda I)X = O$ tiene solución no nula (para la incógnita X); esto último acontece si, y sólo si, la matriz $A - \lambda I$, de dicha ecuación homogénea en la incógnita X, es una matriz singular, como había que comprobar.

2. Un vector $\vec{x} \in V$ es autovector de f asociado al autovalor λ si, y sólo si, $f(\vec{x}) = \lambda\vec{x}$, lo que equivale a que $AX = \lambda X$, lo que significa que $X \in \mathcal{M}_{n\times 1}$ es autovector de A asociado a λ. Obsérvese que la ecuación $AX = \lambda X$ se puede poner en la forma $(A - \lambda I)X = O$, de la que se obtienen, pues, los autovectores de A asociados a λ.

   Como el subespacio propio V_λ (de A) está formado por las soluciones de la ecuación homogénea $(A - \lambda I)X = O$, las cuales forman un subespacio de dimensión n-rang$(A - \lambda I)$ (véase el teorema de Rouché, [106], 2.º), resulta que esta diferencia es la dimensión de V_λ.

EJERCICIO

Sea $f: \mathbb{R}^3 \to R^3$ el endomorfismo definido por

$$(x, y, z) \mapsto (x', y', z') = (3x, x + 2y, 4x + 2z)$$

Hallar los autovalores de f y sus correspondientes subespacios propios.

RESOLUCION

La ecuación $X' = AX$ de f en la base canónica es:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 4 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Por tanto, los autovalores de f son los $\lambda \in \mathbb{R}$ para los que

$$A - \lambda I = \begin{bmatrix} 3-\lambda & 0 & 0 \\ 1 & 2-\lambda & 0 \\ 4 & 0 & 2-\lambda \end{bmatrix} \text{ es singular, que son } \begin{cases} \lambda = 3 \\ \lambda = 2 \end{cases}$$

El subespacio propio V_3, correspondiente a $\lambda = 3$, tiene por ecuación a $(A - 3I)X = O$, esto es:

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 4 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \text{luego } V_3 = \{(\alpha, \alpha, 4\alpha) \,/\, \alpha \in \mathbb{R}\}$$

El subespacio propio V_2, correspondiente a $\lambda = 2$, tiene por ecuación a $(A - 2I)X = O$, esto es:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \text{luego } V_2 = \{(0, \alpha, \beta) \,/\, \alpha, \beta \in \mathbb{R}\}$$

POLINOMIO CARACTERISTICO DE UNA MATRIZ

Según se acaba de decir, en el caso de dimensión finita los autovalores, de un endomorfismo f o de su matriz A, son los escalares λ que hacen de $A - \lambda I$ una matriz singular. Recurriendo a los determinantes, del anterior resultado se desprende que:

[164]

Sea $A = [a_{ij}]$ una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K. Los autovalores de A son los escalares $\lambda \in K$ tales que $\det(A - \lambda I) = 0$, esto es, que:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0 \quad \left(\begin{array}{l} \text{esta ecuación, en la} \\ \text{incógnita } \lambda, \text{ se llama} \\ \textit{ecuación característica} \\ \text{de la matriz } A = [a_{ij}] \end{array} \right)$$

Tomando a λ como indeterminada, $\det(A - \lambda I)$ es un polinomio, con coeficientes en K y de grado n, que se llama *polinomio característico* de la matriz A. Este polinomio es:

$$\det(A - \lambda I) = (-1)^n \lambda^n + (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + \det A \quad (*)$$

Los autovalores de la matriz A son, pues, las raíces $\lambda \in K$ de su polinomio característico.

(*) Sólo se dan aquí los coeficientes de λ^n, λ^{n-1} y λ^0 (término independiente). Respecto de los demás, véase la «observación» que se hace más adelante. Se llama *traza* de A a la suma $\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}$.

COMPROBACION

1. De acuerdo con lo ya obtenido, en el Apartado [163], el escalar $\lambda \in K$ es autovalor de A si, y sólo si, la matriz $A - \lambda I$ es singular, lo que equivale a decir que $\det(A - \lambda I) = 0$, que es la ecuación característica.
2. Desarrollando el determinante de $A - \lambda I$, para lo que los elementos $a_{ii} - \lambda$ de su diagonal se manejan como escalares de K, la expresión a la que se llega es, evidentemente un polinomio de grado n.
3. En el polinomio característico, al desarrollar $\det(A - \lambda I)$, los términos de grado n y los de grado $n - 1$ proceden todos del término de dicho desarrollo que resulta de multiplicar los elementos de su diagonal principal (en los demás términos del desarrollo el mayor grado es $n - 2$); esto es, los términos en λ^n y en λ^{n-1} de $\det(A - \lambda I)$ son los términos en λ^n y en λ^{n-1} de

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) =$$
$$= (-1)^n \lambda^n + (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + (a_{11} a_{22} \cdots a_{nn})$$

que son $(-1)^n$ y $(-1)^{n-1}(a_{11}+a_{22}+\cdots+a_{nn})$, según se había anunciado. Por otra parte, el término independiente, como es el valor de polinomio para $\lambda=0$, valdrá $\det(A-0\lambda)=\det A$, con lo cual concluye la comprobación.

EJERCICIO

Sea A una matriz cuadrada, cuyos elementos pertenecen a un cuerpo K. Compruébese que:

1. Si λ es un autovalor de A, entonces λ^2 es un autovalor de A^2.
2. Si $\lambda \in K$ no es autovalor de A, λ^2 puede ser autovalor de A^2.
3. Puede ocurrir que A^2 tenga algún autovalor y que A carezca de ellos.

RESOLUCION

1. Se sabe que $\det(A-\lambda I)=0$ y hay que probar que $\det(A^2-\lambda^2 I)=0$. Esto es cierto ya que

$$(A+\lambda I)(A-\lambda I)=A^2-\lambda^2 I$$

luego

$$\det(A^2-\lambda^2 I)=[\det(A+\lambda I)][det(A-\lambda I)]=[\det(A+\lambda I)]0=0$$

2. Así ocurre con la siguiente matriz real A, ya que $\lambda=3$ no es autovalor de A y $\lambda^2=9$ si es autovalor de A^2:

$$A=\begin{bmatrix}-3 & 0\\ 1 & -3\end{bmatrix} \quad ; \quad A^2=\begin{bmatrix}9 & 0\\ -6 & 9\end{bmatrix}$$

3. Así ocurre con la siguiente matriz real A, ya que A carece de autovalores (el cuerpo de escalares es $\mathbb{R}$) y $\lambda=-1$ es autovalor de A^2:

$$A=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix} \quad ; \quad A^2=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}$$

OBSERVACION

Anteriormente, al estudiar el polinomio característico $\det(A-\lambda I)$ de una matriz $A=[a_{ij}]$ de tamaño $n\times n$, sólo se obtuvieron tres de sus coeficientes; vamos ahora a determinar todos, comprobando que el término de grado h de dicho polinomio vale

$$(-\lambda)^h\left[\text{Suma de los } \binom{n}{h} \text{ menores diagonales de orden } n-h \text{ de } A\right]$$

donde se han llamado menores diagonales de A a las menores de las submatrices de A que están situadas simétricamente respecto de la diagonal; esto es, aquellas que se obtienen tomando los índices de sus filas iguales a los índices de sus columnas.

Para desarrollar $\det(A - \lambda I)$, descompongamos cada una de las filas de $A - \lambda I$ como suma de las correspondientes filas de A y de $-\lambda I$; con ello, se obtiene el $\det(A - \lambda I)$ como suma de 2^n determinantes, que son, para $h = 1, 2, \ldots, n$, los $\binom{n}{h}$ determinantes que resultan de sustituir h filas de A por las correspondientes filas de $-\lambda I$. El determinante de la matriz que se obtiene de sustituir, en A, las filas $i_1, i_2, \ldots, i_h$ por las correspondientes filas de $-\lambda I$, desarrollándole por estas filas (en cada una de las cuales sólo tiene un elemento no nulo, que vale $-\lambda$) resulta ser igual a $(-\lambda)^h$ por el determinante de la submatriz que se obtiene de prescindir de las filas $i_1, i_2, \ldots, i_h$ y las columnas de los mismos órdenes, que es un menor diagonal de orden $n - h$. Al hacer esto, para cada h fijo, de todos los modos posibles, se obtienen todos los menores diagonales de orden $n - h$, lo que conduce al resultado que se había anunciado.

POLINOMIO CARACTERISTICO DE UN ENDOMORFISMO

[165]

1. Sean A y A' dos matrices cuadradas de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K. Si A y A' son semejantes (es decir, si $A' = P^{-1}AP$ para cierta matriz regular P; véase [094]), entonces A y A' tienen el mismo polinomio característico (y, por ello, la misma ecuación característica):

$$[A' \text{ semejante a } A] \Rightarrow \det(A' - \lambda I) = \det(A - \lambda I)$$

2. Si $f : V \to V$ es un endomorfismo, en el espacio vectorial $V \neq O$ de dimensión n (sea K el cuerpo de escalares), entonces todas las matrices de f (en las distintas bases de V) tienen el mismo polinomio característico y la misma ecuación característica, que se llaman *polinomio característico* y *ecuación característica* del endomorfismo f. Los autovalores de f son aquellos escalares $\lambda \in K$ que son raíces de su polinomio característico.

DEMOSTRACION

1. $A' - \lambda I = P^{-1}AP - \lambda I = P^{-1}AP - P^{-1}\lambda IP = P^{-1}(A - \lambda I)P$, luego

$$\det(A' - \lambda I) = (\det P^{-1})[\det(A - \lambda I)](\det P) =$$
$$= (\det P)^{-1}[\det(A - \lambda I)](\det P) = \det(A - \lambda I)$$

2. Como todas las matrices asociadas a f, en las distintas bases de V, son semejantes entre sí (véase [094]), según lo que se acaba de probar tienen el mismo polinomio característico.

Finalmente, como los autovalores de f son los autovalores de cualquier matriz asociada a f (véase [163], 1), aquellos son los $\lambda \in K$ tales que $\det(A - \lambda I) = 0$ (véase [164]) donde A es cualquier matriz asociada a f, es decir son los $\lambda \in K$ que son raíces del polinomio característico de A, que es el de f.

OBSERVACIONES

1. Sean $A = [a_{ij}]$ y $A' = [a'_{ij}]$ dos matrices cuadradas de igual tamaño. Si A y A' son semejantes, como entonces tienen el mismo polinomio característico, serán iguales los respectivos coeficientes de ambos polinomios y, en particular, A y A' tendrán la misma traza:

$$[A' \text{ semejante a } A] \Rightarrow [a'_{11} + a'_{22} + \cdots + a'_{nn} = a_{11} + a_{22} + \cdots + a_{nn}]$$

2. Como todas las matrices de un mismo endomorfismo (en las distintas bases) son semejantes, todas tienen la misma traza, a la que se llama traza del endomorfismo.

EJERCICIO

Hallar el polinomio característico del endomorfismo $f: \mathbb{R}^5 \to \mathbb{R}^5$ definido mediante:

$$(x_1, x_2, x_3, x_4, x_5) \mapsto (-x_2, x_1 + x_3, 2x_3 - x_4, 2x_4 + 6x_5, 3x_5)$$

RESOLUCION

El polinomio característico de f es el de cualquiera de sus matrices y, en particular, el de su matriz respecto de la base canónica, que es la

$$A = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

El polinomio característico de esta matriz es:

$$\begin{vmatrix} -\lambda & -1 & 0 & 0 & 0 \\ 1 & -\lambda & 1 & 0 & 0 \\ 0 & 0 & 2-\lambda & -1 & 0 \\ 0 & 0 & 0 & 2-\lambda & 6 \\ 0 & 0 & 0 & 0 & 3-\lambda \end{vmatrix} = 0 \quad , \quad \text{esto es: } (3-\lambda)(2-\lambda)^2(\lambda^2+1)$$

9.2. MULTIPLICIDADES ALGEBRAICA Y GEOMETRICA DE UN AUTOVALOR

DEFINICION Y TEOREMA

[166]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$ de dimensión finita n; sea K el cuerpo de escalares. Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a K. Si $\lambda_0 \in K$ es un autovalor de f o de A, se llaman:

- *Multiplicidad algebraica* de λ_0 es el orden de multiplicidad, m, de λ_0 como raíz de la ecuación característica, de f o de A.
- *Multiplicidad geométrica* de λ_0 es la dimensión, d, del subespacio propio, de f o de A, asociado a λ_0.

Si A es matriz de f (en cierta base de V), entonces los autovalores de A y de f son los mismos y cada uno de ellos tiene, en A y en f, la misma multiplicidad algebraica y la misma multiplicidad geométrica.

Si λ_0 es un autovalor, de f o de A, entonces entre su multiplicidad algebraica m y su multiplicidad geométrica d se verifica la relación:

$$1 \leqslant d \leqslant m$$

DEMOSTRACION

Antes de proceder a comprobar que $1 \leqslant d \leqslant m$, conviene recordar que, si A es matriz de f, entonces A y f tienen la misma ecuación característica, pues tienen el mismo polinomio característico (véase [165]), y por ello tienen los mismos autovalores y éstos tienen iguales sus multiplicidades algebraicas, para A y para f. Además, como las dimensiones de los subespacios propios de A y de f asociados a un autovalor λ_0 son, respectivamente:

$$n - \text{rang}\,(A - \lambda_0 I) \quad \text{y} \quad n - \text{rang}\,(f - \lambda_0 i)$$

(I es la matriz unidad $n \times n$ e $i: V \to V$ es la aplicación idéntica) resulta, pues, que éstas son iguales; es decir, las multiplicidades geométricas de λ_0 como autovalor de A y como autovalor de f son iguales.

Si λ_0 es un autovalor, entonces tiene asociado algún autovector no nulo, luego el subespacio propio de λ_0 tiene dimensión igual o mayor que 1, esto es, se verifica que $1 \leqslant d$.

Vamos a suponer que λ_0 es autovalor del endomorfismo f; el caso de λ_0 autovalor de A es equivalente al anterior. Recurramos a una base $(\bar{e}_1, ..., \bar{e}_d)$ del subespacio propio V_0 asociado a λ_0 y completémosla hasta obtener una base $(\bar{e}_1, ..., \bar{e}_d, ..., \bar{e}_n)$ de V. La matriz de f en esta base será, expresada en bloques, del tipo:

$$\left[\begin{array}{c|c} \lambda_0 I & M \\ \hline O & N \end{array}\right] \quad \begin{array}{l} \lambda_0 I \text{ tiene tamaño } d \times d \\ N \text{ es de tamaño } (n-d)\times(n-d) \\ O \text{ matriz nula } (n-d)\times d \end{array}$$

Por tanto, el polinomio característico de f será:

$$\left|\begin{array}{c|c} (\lambda_0 - \lambda)I & M \\ \hline O & N - \lambda I \end{array}\right| = (\lambda_0 - \lambda)^d \det(N - \lambda I) = (\lambda_0 - \lambda)^d p(\lambda)$$

donde $p(\lambda) = \det(N - \lambda I)$ es un polinomio de grado $n - d$. Por tanto, λ_0 es raíz del polinomio característico de f con orden de multiplicidad mayor o igual que d, esto es, se verifica que $m \geq d$ (es de orden d si no es raíz de $p(\lambda)$; es de orden mayor que d si es raíz de $p(\lambda)$).

EJEMPLO

Consideremos la siguiente matriz A

$$A = \begin{bmatrix} 1 & 2 & c \\ 0 & 1 & 0 \\ 0 & 4 & 3 \end{bmatrix} \quad (c \in \mathbb{R} \text{ dado})$$

Recurriendo a su ecuación característica, resulta que sus autovalores son:

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 & c \\ 0 & 1-\lambda & 0 \\ 0 & 4 & 3-\lambda \end{vmatrix} = (1-\lambda)^2(3-\lambda) = 0 \rightarrow \begin{cases} \lambda_1 = 3 \text{ (simple)} \\ \lambda_2 = 1 \text{ (doble)} \end{cases}$$

El subespacio propio correspondiente a $\lambda_1 = 3$ es:

$$V_{\lambda=3}: \begin{bmatrix} -2 & 2 & c \\ 0 & -2 & 0 \\ 0 & 4 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow V_{\lambda=3} = \{(c\alpha, 0, 2\alpha \;/\; \alpha \in \mathbb{R}\}$$

las multiplicidades algebraica y geométrica de $\lambda_1 = 3$ son pues $m = 1$ y $d = 1$.
El subespacio propio correspondiente a $\lambda_2 = 1$ es:

$$V_{\lambda=1}: \begin{bmatrix} 0 & 2 & c \\ 0 & 0 & 0 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{cases} \text{Si } c = 1,\ V_{\lambda=1} = \{(\alpha, \beta, -2\beta) \;/\; \alpha, \beta \in \mathbb{R}\} \\ \text{Si } c \neq 1,\ V_{\lambda=1} = \{(\alpha, 0, 0) \;/\; \alpha, \beta \in \mathbb{R}\} \end{cases}$$

La multiplicidad algebraica de $\lambda_2 = 1$ es $m = 2$ (para cualquiera que sea c); la multiplicidad geométrica de $\lambda_2 = 1$ es $d = 2$ ó $d = 1$ según que sea $c = 1$ ó $c \neq 1$, respectivamente.

OBSERVACIONES

1. Un «autovalor simple», de un endomorfismo o de una matriz, es un escalar $\lambda_0 \in K$ (donde K es el cuerpo de escalares) que es raíz simple de la ecuación característica, del endomorfismo o de la matriz; es decir, es raíz con orden de multiplicidad $m = 1$. Como la multiplicidad geométrica d de λ_0 es tal que $1 \leq d \leq m$, resulta que ha de ser $d = 1$. El subespacio propio asociado a un autovalor simple tiene, pues, dimensión 1.
2. Dos matrices cuadradas A y A' que sean semejantes tienen los mismos autovalores y cada uno de éstos tiene, respecto de A y respecto de A', la misma multiplicidad algebraica y la misma multiplicidad geométrica. Esto se desprende del hecho de estar A y A' asociadas a un mismo endomorfismo.
3. La recíproca de la anterior propiedad es falsa. Es decir, dos matrices cuadradas de igual tamaño, A y A', que tenga los mismos autovalores y que cada uno de éstos tenga, respecto de A y de A', la misma multiplicidad algebraica y la misma multiplicidad geométrica pueden, sin embargo no ser matrices semejantes. Así ocurre, por ejemplo, con las siguientes matrices A y A' reales:

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad ; \quad A' = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Ambas tienen un solo autovalor, que es $\lambda = 2$, y éste tiene, en A y en A', multiplicidad algebraica $m = 4$ y multiplicidad geométrica $d = 2$. No obstante, A y A' no son semejantes, ya que la relación $A' = P^{-1}AP$, donde P es matriz regular, o su equivalente $PA' = AP$ no tiene solución (con P regular). Esto se comprueba fácilmente tomando una matriz genérica P, de tamaño 4×4 (con 16 parámetros), realizando los productos PA' y AP e identificando; haciendo esto, nos encontramos con que P habría de ser (para ciertos números reales a, b, c, d, e y f):

$$\begin{bmatrix} a & b & c & d \\ 0 & a & 0 & c \\ 0 & 0 & 0 & 0 \\ 0 & e & 0 & f \end{bmatrix}$$

que no puede ser regular (tiene su tercera fila nula).

NUMERO MAXIMO DE AUTOVECTORES INDEPENDIENTES

[167]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$ de dimensión finita n; llamamos K al cuerpo de los escalares. Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen al cuerpo K.

Si los autovalores distintos, que tiene f o A, son $\lambda_1, \lambda_2, ..., \lambda_p \in K$ y si las multiplicidades geométricas de éstos se denotan por $d_i = \dim V_{\lambda_i}$ para $i = 1, 2, ..., p$, entonces el número máximo de vectores propios linealmente independientes, de f o de A, es $d_1 + d_2 + \cdots + d_p$.

Nótese que, si m_i es la multiplicidad algebraica de λ_i, se verifica:

$$p \leq d_1 + d_2 + \cdots + d_p \leq m_1 + m_2 + \cdots + m_p \leq n.$$

COMPROBACION

Como los subespacios propios V_{λ_i} (formados por los vectores propios) son independientes, esto es, tienen suma directa (véase [162]), el número máximo de vectores propios linealmente independientes se consigue tomando una base en cada uno de ellos y uniendo todas ellas. Puesto que una base de V_{λ_i} consta de d_i vectores, resulta que el número máximo de vectores propios, de f o de A, linealmente independientes es $d_1 + d_2 + \cdots + d_p$.

La relación que figura al final del anterior enunciado es consecuencia obvia de las relaciones $1 \leq d_i \leq m_i$ para $i = 1, 2, ..., p$ (véase [166]).

RECORDATORIO

En lo que sigue, vamos a precisar de algunas cuestiones relativas a las raíces de los polinomios y a las multiplicidades de aquéllas. Comencemos recordando estos resultados:

[168]

1. Sea $p(x)$ un polinomio, con una indeterminada x, cuyos coeficientes son elementos de un cuerpo K (interesan especialmente los casos $K = \mathbb{R}$ y $K = \mathbb{C}$). Si $\lambda \in K$ es tal que $p(\lambda) = 0$, se dice que λ es una *raíz* de $p(x)$. Si $p(\lambda) = 0$ para un λ perteneciente a un cuerpo H más amplio que K, se dice que λ es una raíz de $p(x)$ en H. Para que λ sea una raíz de $p(x)$ es necesario y suficiente que $p(x)$ sea divisible por $x - \lambda$, esto es, que $p(x) = (x - \lambda)q(x)$, donde $q(x)$ es otro polinomio.

2. Se llama *multiplicidad* u orden de multiplicidad de una raíz λ del polinomio $p(x)$ al mayor de los números naturales m tales que $p(x)$ sea divisible por $(x-\lambda)^m$; si $m = 1, 2, 3$, etc.. se dice que λ es raíz simple, doble, triple, etc., de $p(x)$. Si $p(x)$ tiene grado n, si las raíces de $p(x)$ son $\lambda_1, \lambda_2, ..., \lambda_p$ y si las multiplicidades de éstas son, respectivamente, $m_1, m_2, ..., m_p$, entonces se verifica que, para cierto polinomio $q(x)$ (que puede ser constante), es:

$$p(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_p)^{m_p}q(x) \qquad [1]$$

donde

$$m_1 + m_2 + \cdots + m_p \leqslant n \quad (n = \text{grado de } p(x))$$

(si el grado de $p(x)$ es n, entonces $p(x)$ tiene a lo sumo n raíces).

3. Un polinomio complejo (esto es, que tiene sus coeficientes complejos) $p(x)$ de grado n tiene exactamente n raíces complejas, si se cuenta cada una de ellas tantas veces como indica su multiplicidad; esto es, la anterior descomposición [1] es ahora, que $K = \mathbb{C}$:

$$p(x) = c(x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_p)^{m_p}$$

con

$$m_1 + m_2 + \cdots + m_p = n \quad (\text{grado de } p(x))$$

$$c \in \mathbb{C} \text{ coeficiente del término de mayor grado}$$

Cuando con un cuerpo K ocurre, como con $\mathbb{C}$, que todo polinomio (de grado $n \geqslant 1$) con coeficientes en K tiene alguna raíz en K, se dice que K es *algebraicamente cerrado*.

4. El cuerpo $\mathbb{R}$ no es algebraicamente cerrado. La descomposición de un polinomio real $p(x)$ en producto polinomios irreducibles reales tiene, en general, polinomios de primer y de segundo grado, esto es de los tipos:

$$x - \lambda \quad \text{y} \quad x^2 + rx + s$$

con

$$r^2 - 4s < 0$$

Pero todo polinomio $p(x)$ de grado $n \geqslant 1$ con coeficientes reales tiene n raíces complejas (algunas de ellas, o todas, pueden ser reales), si se cuenta cada una tantas veces como indica su multiplicidad. Si λ es una raíz compleja del polinomio real $p(x)$, entonces $\bar{\lambda}$ (número complejo conjugado de λ) también es raíz de $p(x)$ y lo es con la misma multiplicidad que λ.

MULTIPLICIDADES ALGEBRAICAS EN LOS CASOS REAL Y COMPLEJO

[169]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial real o complejo $V \neq O$, de dimensión finita n. Sea también A una matriz cuadrada real o compleja de tamaño $n \times n$.

1. En el caso complejo, el endomorfismo f (complejo) y la matriz A (compleja) tienen al menos un autovalor (complejo(*)); si cada autovalor se cuenta un número de veces igual a su multiplicidad algebraica, entonces f y A tienen exactamente n autovalores (complejos(*)).
2. En el caso real, el endomorfismo f (real) o la matriz A (real) pueden carecer de autovalores (reales) y, caso de tenerlos, a lo sumo tienen n autovalores (reales), contando cada uno tantas veces como indique su multiplicidad algebraica. No obstante, f (real) y A (real) tienen siempre n autovalores complejos (contando cada uno el número de veces que señala su multiplicidad algebraica).

(*) Cuando aquí decimos que un autovalor es complejo, no se descarta la posibilidad de que, en particular, pudiera ser real; esto es, que tuviera nula su parte imaginaria.

DEMOSTRACION

Tanto para f como para A, sus autovalores son las raíces del polinomio característico (véase [164]), que tiene grado n. Por tanto, las anteriores afirmaciones 1 y 2 son reflejo de lo dicho en [168] 3 y 4, respectivamente.

9.3. COMPLEJIFICACION DE UN ESPACIO VECTORIAL REAL

En el estudio de los endomorfismos de un espacio vectorial, muchos resultados son válidos tanto si los escalares son números reales como si son números complejos. Pero no ocurre así siempre. Una causa de ello, que se ha hecho ya notar en el anterior resultado, es el hecho de que $\mathbb{C}$ sea un cuerpo algebraicamente cerrado, cuando $\mathbb{R}$ no lo es. También conviene notar que hay propiedades, que se verifican en ambos casos, real y complejo, pero que se prueban más fácilmente en el segundo de ellos.

Por esto, dado un espacio vectorial real V, se le va a asociar un espacio vectorial complejo V_C que goce de las propiedades de V, al objeto de estudiar éstas en V_C y, luego, «trasladarlas» a V.

PROPOSICION

[170]

Sea $V \neq O$ un espacio vectorial real. Considérese el producto cartesiano $V \times V = V^*$. Este conjunto es un espacio vectorial complejo respecto de las siguientes operaciones suma y producto por un número complejo $a + bi$:

$$(\bar{u}, \bar{v}) + (\bar{u}', \bar{v}') = (\bar{u} + \bar{u}', \bar{v} + \bar{v}'),$$

$$(a + bi)(\bar{u}, \bar{v}) = (a\bar{u} - b\bar{v}, b\bar{u} + a\bar{v}); \qquad (1)$$

a este espacio vectorial le denotaremos por V_C^*. Si sólo se consideran escalares reales (o sea, si $b = 0$), este espacio vectorial pasa a ser un espacio vectorial real, al que denotaremos por V_R^*.

El conjunto $V \times O = \{(\bar{u}, \bar{o}) \,/\, \bar{u} \in V\}$ es un subespacio vectorial de V_R^* isomorfo a V, con el que se identifica poniendo $(\bar{u}, \bar{o}) = \bar{u}$. El conjunto $O \times V = \{(\bar{o}, \bar{v}) \,/\, \bar{v} \in V\}$ es un subespacio vectorial de V_R^* isomorfo a V, al que se denota por iV, ya que $(\bar{o}, \bar{v}) = i(\bar{v}, \bar{o}) = i\bar{v}$. Los subespacios $V \times O = V$ y $O \times V = iV$ son suplementarios de V_R^*. Se verifica (tanto para V_R^* como para V_C^*) que:

$$V^* = V \oplus iV = \{\bar{u} + i\bar{v} \,/\, \bar{u}, \bar{v} \in V\}$$

El espacio vectorial V_C^* se llama la *complejificación* del espacio vectorial real V.

---

(1) Esta operación se define como se haría si $a + b_i$ y $(\bar{u}, \bar{v})$ fuesen los números complejos $(a, b) = a + b_i$ y $(u, v) = u + iv$.

COMPROBACIONES

- La axiomática de espacio (véase [052]) se verifica para V_C^*, como se comprueba trivialmente. Nótese que el elemento nulo de V_C^* es el $(\bar{o}, \bar{o})$ y que el opuesto de $(\bar{u}, \bar{v})$ es el $(-\bar{u}, -\bar{v})$.
- Aún es más sencillo comprobar que V_R^* es un espacio vectorial; no se pierda de vista que, aquí, el producto por un escalar $a \in \mathbb{R}$ viene dado por

$$a(\bar{u}, \bar{v}) = (a\bar{u}, a\bar{v})$$

- La identificación $(\bar{u}, \bar{o}) = \bar{u}$, para $\bar{u} \in V$, es lícita ya que la aplicación

$$V \times O \to V, \quad (\bar{u}, \bar{o}) \mapsto \bar{u}$$

es un isomorfismo del espacio vectorial $V \times O$ (subespacio de V_R^*) en el V. El conjunto $O \times V$ es también un subespacio de V_R^* y

$$O \times V \to V, \quad (\bar{o}, \bar{v}) \mapsto \bar{v}$$

es un isomorfismo de espacios vectoriales, como se comprueba fácilmente. Conviene observar que, de acuerdo con la definicioín de producto por un número complejo, y como se puede poner $\bar{v} = (\bar{v}, \bar{o})$, es:

$$i\bar{v} = i\,(\bar{v}, \bar{o}) = (\bar{o} - \bar{o}, \bar{v} + \bar{o}) = (\bar{o}, \bar{v})$$

luego

$$O \times V = \{(\bar{o}, \bar{v}) \,/\, \bar{v} \in V\} = \{i\bar{v} \,/\, \bar{v} \in V\}$$

lo que conduce a denotar $O \times V$ como iV.

- Nótese linealmente que todo $(\bar{u}, \bar{v}) \in V^2$ se puede poner como suma en la forma:

$$(\bar{u}, \bar{v}) = (\bar{u}, \bar{o}) + (\bar{o}, \bar{v}) = \bar{u} + i\bar{v}$$

luego $V + iV = V^2$. Como además $V \cap iV = O$, resulta que V e iV son subespacios suplementarios de V_R^*. Nótese que la descomposición $(\bar{u}, \bar{v}) = \bar{u} + i\bar{v}$ sigue siendo válida y única en V_C^*.

PROPOSICION

[171]

Sea $V \neq O$ un espacio vectorial real y sea V_C^* la complejificación de V (véase [170]). Se verifica que:

1. Si $(\bar{e}_1, ..., \bar{e}_n)$ es un sistema linealmente independiente de vectores de V, o si es un generador de V o si es una base de V, entonces $(\bar{e}_1, ..., \bar{e}_n)$ también es, respectivamente, linealmente independiente en V_C^*, un generador de V_C^* o una base de V_C^*.
2. Si $f: V \to V$ es un endomorfismo, entonces

$$f_C: V_C^* \to V_C^*, \quad f_C(\bar{u} + i\bar{v}) = f(\bar{u}) + if(\bar{v})$$

es un endomorfismo en V_C^*. La restricción de f_C a V es f. Además, si V tiene dimensión finita y A es la matriz de f en cierta base de V, entonces A también es la matriz de f_C, en la misma base.
3. Sea $\lambda_0 \in \mathbb{C}$ un autovalor del anterior endomorfismo $f_C: V_C^* \to V_C^*$. Si λ_0 es real, entonces λ_0 es un autovalor del endomorfismo $f: V \to V$. Si $\lambda_0 = a + bi$ es complejo (no real) y $\bar{u} + i\bar{v}$ es vector propio de f_C correspondiente a λ_0, entonces $V_0 = \mathcal{V}(\bar{u}, \bar{v})$ es un subespacio, de V, invariante por f, es decir, tal que $f(V_0) \subset V_0$; en concreto $f(\bar{u}) = a\bar{u} - b\bar{v}$ y $f(\bar{v}) = b\bar{u} + a\bar{v}$. Además $\bar{\lambda}_0 = a - bi$ es también un autovalor de f_C y $\bar{u} - i\bar{v}$ es vector propio de f_C asociado al autovalor $\bar{\lambda}_0$.

DEMOSTRACION

1. Veamos que son ciertas las tres proposiciones relativas a $S = (\bar{e}_1, ..., \bar{e}_n)$:

   • Si S es linealmente independiente en V, también lo es en V_C^* ya que, como $V^* = V \oplus iV$, será:

$$(a_1 + ib_1)\bar{e}_1 + \cdots + (a_n + ib_n)\bar{e}_n = \bar{o} \Rightarrow$$

$$\Rightarrow (a_1\bar{e}_1 + \cdots + a_n\bar{e}_n) + i(b_1\bar{e}_1 + \cdots + b_n\bar{e}_n) = \bar{o} \Rightarrow$$

$$\Rightarrow \left.\begin{matrix} a_1\bar{e}_1 + \cdots + a_n\bar{e}_n = \bar{o} \Rightarrow a_1 = \cdots = a_n = 0 \\ b_1\bar{e}_1 + \cdots + b_n\bar{e}_n = \bar{o} \Rightarrow b_1 = \cdots = b_n = 0 \end{matrix}\right\} \Rightarrow a_j + ib_j = 0 \quad \text{para } j = 1, ..., n$$

   • Si S genera V, entonces también genera V_C^*, ya que dado $(\bar{u}, \bar{v}) \in V_C^*$, como $\bar{u}, \bar{v} \in V$, existen $a_1, ..., a_n \in \mathbb{R}$ y $b_1, ..., b_n \in \mathbb{R}$ tales que (para $j = 1, 2, ..., n$):

$$\left.\begin{matrix} \bar{u} = \sum_j a_j\bar{e}_j \\ \bar{v} = \sum_j b_j\bar{e}_j \end{matrix}\right\} \text{ luego } (\bar{u}, \bar{v}) = \bar{u} + i\bar{v} = \sum_j (a_j + ib_j)\bar{e}_j$$

   • Si S es base de V, entonces es independiente en V y genera V; entonces, según las dos propiedades anteriores, S es independiente en V_C^* y genera V_C^*, o sea, es base de V_C^*.

2. La linealidad de f_C es evidente, ya que para $\lambda, \lambda' \in \mathbb{C}$ y para $\bar{u} + i\bar{v}$ y $\bar{u}' + i\bar{v}'$ de V_C^*, es:

$$f_C[\lambda(\bar{u} + i\bar{v}) + \lambda'(\bar{u}' + i\bar{v}')] = f_C[(\lambda\bar{u} + \lambda'\bar{u}') + i(\lambda\bar{v} + \lambda'\bar{v}')] = f(\lambda\bar{u} + \lambda'u') +$$

$$+ if(\lambda\bar{v} + \lambda'\bar{v}') = \lambda f(\bar{u}) + \lambda' f(\bar{u}') + i[\lambda f(\bar{v}) + \lambda' f(\bar{v}')] =$$

$$= \lambda[f(\bar{u}) + if(\bar{v})] + \lambda'[f(\bar{u}') + if(\bar{v}')] = \lambda f_C(\bar{u} + i\bar{v}) + \lambda' f_C(\bar{u}' + i\bar{v}')$$

   La restricción de f_C a V es f ya que para cualquier $\bar{u} \in V$ es:

$$f_C(\bar{u}) = f_C(\bar{u} + i\bar{o}) = f(\bar{u}) + if(\bar{o}) = f(\bar{u}) + i\bar{o} = f(\bar{u})$$

   Si la base que se toma en V es $(\bar{e}_1, ..., \bar{e}_n)$, las columnas de la matriz A de f son las columnas de coordenadas de $f(\bar{e}_1), ..., f(\bar{e}_n)$. Llamando A_C a la matriz de f_C en la misma base, las columnas de A_C son las columnas de coordenadas de $f_C(\bar{e}_1), ..., f_C(\bar{e}_n)$, que son las mismas que las de $f(\bar{e}_1), ..., f(\bar{e}_n)$, ya que la restricción de f_C a V es f; como A y A_C tienen las mismas columnas, son iguales.

3. Como $\lambda_0 \in \mathbb{C}$ es un autovalor de f_C, existe un vector no nulo $\bar{u} + i\bar{v} \in V^*$ tal que

$$f_C(\bar{u} + i\bar{v}) = \lambda_0(\bar{u} + i\bar{v}), \quad \text{luego } f(\bar{u}) + if(\bar{v}) = \lambda_0\bar{u} + \lambda_0 i\bar{v}$$

- Si λ_0 es real, entonces $f(\bar{u})$ y $\lambda_0\bar{u}$ son de V e $if(\bar{v})$ y $\lambda_0 i\bar{v}$ son de iV; por tanto, como la suma $V + iV$ es directa, de la última igualdad se deduce que:

$$f(\bar{u}) = \lambda_0\bar{u} \quad \text{y} \quad f(\bar{v}) = \lambda_0\bar{v} \qquad [1]$$

Ahora bien, como $\bar{u} + i\bar{v} \neq \bar{o}$, al menos uno de los $\bar{u}$ o $\bar{v}$ es no nulo y, por ello, las relaciones [1] permiten decir que $\lambda_0 \in \mathbb{R}$ es un autovalor de f.

- Si λ_0 es complejo, o sea, si $\lambda_0 = a + bi$ con $b \neq 0$ $(a, b \in \mathbb{R})$, y si $\bar{u} + i\bar{v}$ es un vector propio de f_C asociado a λ_0, será:

$$\left.\begin{aligned} f_C(\bar{u} + i\bar{v}) &= (a + bi)(\bar{u} + i\bar{v}) = (a\bar{u} - b\bar{v}) + i(b\bar{u} + a\bar{v}) \\ f_C(\bar{u} + i\bar{v}) &= f(\bar{u}) + if(\bar{v}) \end{aligned}\right\}$$

luego

$$\begin{cases} f(\bar{u}) = a\bar{u} - b\bar{v} \\ f(\bar{v}) = b\bar{u} + a\bar{v} \end{cases}$$

Estas dos igualdades prueban que f transforma a los vectores de $V_0 = \mathcal{V}(\bar{u}, \bar{v})$ en vectores de V_0.

Finalmente, para comprobar que $\bar{\lambda}_0$ y $\bar{u} - i\bar{v}$ son autovalor y autovector de f_C, se tiene:

$$f_C(\bar{u} - i\bar{v}) = f_C(\bar{u} + i(-v)) = f(\bar{u}) + if(-\bar{v}) = f(u) - if(\bar{v}) =$$

$$= (a\bar{u} - b\bar{v}) - i(b\bar{u} + a\bar{v}) = (a - bi)(\bar{u} - i\bar{v}) = \bar{\lambda}_0(\bar{u} - i\bar{v})$$

Obsérvese que si V tiene dimensión finita, tomando una base cualquiera en V (que lo será también de V^*), los endomorfismos f_C y f tienen la misma matriz A y, por ello, también tienen el mismo polinomio característico $\det(A - \lambda I) = 0$. Las raíces en $\mathbb{R}$ de este polinomio son los autovalores de f; las raíces en $\mathbb{C}$ de este polinomio son los autovalores de f_C.

β ENDOMORFISMOS DIAGONALIZABLES

9.4. DIAGONALIZACION (POR SEMEJANZA)

Aunque algo ya se ha dicho sobre lo que es diagonalizar un endomorfismo o diagonalizar por semejanza la matriz de éste, se hace necesario que concretemos estos conceptos:

[172]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$ de dimensión finita n; sea K el cuerpo de escalares. Se dice que f es *diagonalizable* si existe una base de V en la que la matriz de f es diagonal. Diagonalizar f es hallar aquella base y esta matriz diagonal.

Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K. Se dice que A es *diagonalizable por semejanza* si existe una matriz diagonal semejante a A; esto es, si existe una matriz regular P y una matriz diagonal D tales que $D = P^{-1}AP$. Diagonalizar A por semejanza es hallar las anteriores matrices P y D.

Si A es matriz de f, en cierta base de V, entonces f es diagonalizable si y sólo si A es diagonalizable por semejanza. Si $D = P^{-1}AP$ es diagonalización de A, entonces D es la matriz de f en una nueva base y P es la matriz del correspondiente cambio de coordenadas.

COMPROBACION

- Si f es diagonalizable, llamando P a la matriz del cambio de coordenadas que se produce al pasar a una base en la que f tiene matriz diagonal, esta matriz será $P^{-1}AP$ (véase [094]), luego A es diagonalizable por semejanza.
- Si A es diagonalizable por semejanza, existe una matriz regular P tal que $P^{-1}AP$ es diagonal. Entonces P es matriz de un cambio de coordenadas tal que en la nueva base f tiene por matriz (véase [094]) a la $P^{-1}AP$, que es diagonal, luego f es diagonalizable.

EJERCICIO

Acudiendo a la anterior definición (único recurso de que disponemos, de momento), compruébese que el endomorfismo

$$f: \mathbb{R}^2 \to \mathbb{R}^2, \qquad (x, y) \mapsto (x + 2y, y)$$

no es diagonalizable.

RESOLUCION

La matriz de f en la base canónica de $\mathbb{R}^2$ es:

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Hemos de comprobar que no puede verificarse que $P^{-1}AP = D$, donde P es matriz regular y D es matriz diagonal. De verificarse esto último, también sería

$AP = PD$, que expresada en función de los elementos (desconocidos) de P y de D será:

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x & y \\ z & t \end{bmatrix} = \begin{bmatrix} x & y \\ z & t \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

Esta igualdad equivale a las

$$x + 2z = ax \quad , \quad y + 2t = by \quad , \quad z = az \quad , \quad t = bt$$

que conducen a $a = 1$, $b = 1$, $z = 0$, $t = 0$. Por tanto, la segunda fila de P habría de ser nula, luego P sería singular, lo que prueba la no diagonalización de f.

CARACTERIZACION DE LOS ENDOMORFISMOS DIAGONALIZABLES

Recurriendo a lo hasta aquí dicho, en este capítulo, acerca de los autovalores y de los vectores propios de los endomorfismos y de las matrices, el estudio de su diagonalizabilidad no ofrece ya dificultades:

[173]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$ de dimensión finita n; sea K el cuerpo de escalares. Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen al cuerpo K.

1. El endomorfismo f, o la matriz A, es diagonalizable si y sólo si existe una base de V, o de $\mathcal{M}_{n\times 1}$, respectivamente, formada por vectores propios de f, o de A.
2. Si los autovalores distintos de f, o de A, son $\lambda_1, \lambda_2, ..., \lambda_p \in K$, si las multiplicidades algebraicas de éstos son $m_1, m_2, ..., m_p$ y si sus multiplicidades geométricas son $d_1, d_2, ..., d_p$, entonces f, o A, es diagonalizable si y sólo si se verifican las dos condiciones siguientes:

   - $m_1 + m_2 + \cdots + m_p = n$ (*).
   - $d_i = m_i$ para $i = 1, 2, ..., p$.

   En particular si f, o A, tiene n autovalores distintos (n raíces distintas de su polinomio característico) en K, entonces f, o A, es diagonalizable.

(*) Si $K = \mathbb{C}$, esta condición siempre se verifica.

DEMOSTRACION

Consideraremos el caso del endomorfismo f; el caso de la matriz A es consecuencia obvia del anterior.

1. Supongamos primero que existe una base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V formada por vectores propios de f. Se sabe, pues, que $f(\bar{e}_i) = \lambda_i \bar{e}_i$ $(i = 1, 2, ..., n)$ para ciertos escalares λ_i (autovalores); es decir, $f(\bar{e}_i)$ tiene todas sus coordenadas nulas excepto la de lugar i que vale λ_i. La matriz de f en esta base tiene por columna i-ésima $(i = 1, 2, ..., n)$ a la columna de coordenadas de $f(\bar{e}_i)$; por tanto, de lo que acabamos de decir se deduce que la matriz de f en la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es la matriz diagonal:

$$D = \begin{bmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{bmatrix} \qquad [1]$$

esto es, f es diagonalizable.

Recíprocamente, supongamos que f es diagonalizable; esto es que existe una base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ en la que la matriz de f es diagonal; a esta matriz diagonal la denotaremos como en [1]. Resulta evidente que:

$$f(\bar{e}_1) = \lambda_1 \bar{e}_1, \quad f(\bar{e}_2) = \lambda_2 \bar{e}_2, ..., f(\bar{e}_n) = \lambda_n \bar{e}_n$$

lo que prueba que $\bar{e}_1, \bar{e}_2, ..., \bar{e}_n$ son vectores propios. Así pues, la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ está formada por vectores propios.

2. Como el número máximo de vectores propios de f linealmente independientes es $d_1 + d_2 + \cdots + d_p$ (véase [167]), para que exista una base formada por vectores propios de f es necesario y suficiente que sea

$$d_1 + d_2 + \cdots + d_p = n \qquad [2]$$

lo que caracteriza, pues, a la diagonalizabilidad de f.

Por otra parte, sabemos que (véase [166])

$$d_i \leq m_i (i = 1, 2, ..., p) \quad \text{y} \quad m_1 + m_2 + \cdots m_p \leq n \qquad [3]$$

A la vista de las relaciones [3], es evidente que la condición [2] es equivalente a las:

$$d_i = m_i \ (i = 1, 2, ..., p) \quad \text{y} \quad m_1 + m_2 + \cdots + m_p = n \qquad [4]$$

las cuales son, en consecuencia, condiciones necesarias y suficientes para que f sea diagonalizable.

Finalmente, si hay n autovalores en K, entonces es

$$p = n, \quad d_i = m_i = 1 \ (i = 1, 2, \ldots n) \quad \text{y} \quad m_1 + m_2 + \cdots + m_n = n$$

luego se verifican las dos condiciones [4], lo que garantiza que f, o A, es diagonalizable.

EJERCICIO

Sea $f: \mathbb{R}^4 \to \mathbb{R}^4$ el endomorfismo definido mediante

$$f(x_1, x_2, x_3, x_4) = (x_1 + ax_2, x_2, ax_3 + x_4, -x_3 - ax_4)$$

donde a es un número real dado. Analícese si f es diagonalizable para algún valor de $a \in \mathbb{R}$.

RESOLUCION

La matriz de f en la base canónica de $\mathbb{R}^4$ es la siguiente matriz A, cuya ecuación característica y cuyos autovalores también se dan:

$$A = \begin{bmatrix} 1 & a & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & -1 & -a \end{bmatrix} \qquad \det(A - \lambda I) \equiv (a - \lambda)^2(\lambda^2 - a^2 + 1) = 0$$

$$\begin{cases} \lambda_1 = a \ \text{(doble)} \\ \lambda_2, \ \lambda_3 = \pm\sqrt{a^2 - 1} \ \text{(si } |a| \geq 1) \end{cases}$$

Si $|a| < 1$, entonces sólo hay una raíz real doble, luego f no será diagonalizable pues $\Sigma m_i = 2 \neq 4$. Si $|a| = 1$ y si $|a| > 1$ entonces

$$\Sigma m_i = 2 + 2 = 4 \quad \text{y} \quad \Sigma m_i = 2 + 1 + 1 = 4$$

respectivamente; por ello, en lo que a Σm_i se refiere, f pudiera ser diagonalizable si $|a| \geq 1$. Por otra parte, el subespacio propio correspondiente a λ_1 es:

$$V_{\lambda_1} = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \ / \ ax_2 = 0, \ x_4 = 0, \ -x_3 - 2ax_4 = 0\}$$

cuya dimensión es $d_1 = 2$ si $a = 0$ y es $d_1 = 1$ en los demás casos; para que f sea diagonalizable se precisa pues que sea $a = 0$. Como ya se vio que también se necesitaba que fuese $|a| \geq 1$, resulta que f no es diagonalizable para ningún valor real del parámetro a.

FORMA DIAGONAL

[174]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$ de dimensión finita n; sea K el cuerpo de escalares. Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen al cuerpo K. Suponiendo que f es diagonalizable y que A es diagonalizable por semejanza, las formas diagonales de f y A son:

- La matriz de f es diagonal en una base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V formada uniendo bases de los subespacios propios de f; dicha matriz diagonal es la:

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & O & \\ & O & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

($\lambda_1, \lambda_2, ..., \lambda_n$ son los autovalores de f, cada uno de ellos repetido tantas veces como indique su multiplicidad algebraica; λ_i es el autovalor asociado al vector $\bar{e}_i$)

- La forma diagonal de A, esto es, la anterior matriz D (λ_i son ahora los autovalores de A), es igual a $D = P^{-1}AP$, donde las columnas de P son columnas de coordenadas de los vectores de las bases de los respectivos subespacios propios de A.

COMPROBACION

Aquí no se ha hecho otra cosa que recoger resultados que aparecen en la última demostración realizada.

EJEMPLO

Veamos que el siguiente endomorfismo $f: \mathbb{R}^3 \to \mathbb{R}^3$ es diagonalizable y obtengamos su diagonalización:

$$f(x, y, z) = (x + 2y + 10z,\ 2x + y + 10z,\ -x - y - 6z)$$

RESOLUCION

Hallemos los autovalores de f y sus multiplicidades, para lo que recurriremos a la matriz A de f en la base canónica y al polinomio característico de ésta:

$$A = \begin{bmatrix} 1 & 2 & 10 \\ 2 & 1 & 10 \\ -1 & -1 & -6 \end{bmatrix} \text{ y } \begin{vmatrix} 1-\lambda & 2 & 10 \\ 2 & 1-\lambda & 10 \\ -1 & -1 & -6-\lambda \end{vmatrix} = -\lambda^3 - 4\lambda^2 - 5\lambda - 2$$

Los autovalores de A, esto es, las soluciones de la ecuación $-\lambda^3-4\lambda^2-5\lambda-2=0$ son $\lambda_1=-1$ doble ($m_1=2$) y $\lambda_2=-2$ simple ($m_2=1$). Hallemos los subespacios propios, $V_{\lambda=-1}$ y $V_{\lambda=-2}$, de A:

$$V_{\lambda=-1}:(A+I)X=O\rightarrow\begin{bmatrix}2&2&10\\2&2&10\\-1&-1&-5\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

luego

$$V_{\lambda=-1}=\{(x,y,z)\in\mathbb{R}^3\,/\,x+y+5z=0\}=\mathcal{V}((1,-1,0),(5,0,-1))$$

$$V_{\lambda=-2}:(A+2I)X=O\rightarrow\begin{bmatrix}3&2&10\\2&3&10\\-1&-1&-4\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

luego

$$V_{\lambda=-2}=\{(x,y,z)\in\mathbb{R}^3\,/\,3x+2y+10z=0,\ 2x+3y+10z=0\}=\mathcal{V}(2,2,-1)$$

Como $m_1+m_2=3=\dim\mathbb{R}^3$ y como las dimensiones de $V_{\lambda=-1}$ y $V_{\lambda=-2}$ son $d_1=2=m_1$ y $d_2=1=m_2$, resulta que f es diagonalizable. La diagonalización de f se obtiene en la base $(\bar{u}_1,\bar{u}_2,\bar{u}_3)$:

$$\bar{u}_1=(1,-1,0)\quad,\quad\bar{u}_2=(5,0,-1),\quad\bar{u}_3=(2,2,-1)$$

La matriz de f en esta base es la:

$$D=\begin{bmatrix}-1&0&0\\0&-1&0\\0&0&-2\end{bmatrix}$$

Nótese que $D=P^{-1}AP$, siendo P la matriz cuyas columnas son las columnas de coordenadas de $\bar{u}_1$, $\bar{u}_2$ y $\bar{u}_3$, es decir:

$$P=\begin{bmatrix}1&5&2\\-1&0&2\\0&-1&-1\end{bmatrix}$$

EJERCICIO

Comprobar que la siguiente matriz real A no es diagonalizable, como matriz real. Compruébese que al considerar a A como matriz compleja, entonces sí es diagonalizable:

$$A=\begin{bmatrix}1&i&0&1+i\\-i&1&0&1-i\\0&-i&i&1\\1+i&-1+i&0&2i\end{bmatrix}$$

RESOLUCION

El polinomio característico de A es:

$$\begin{vmatrix} 1-\lambda & i & 0 & 1+i \\ -i & 1-\lambda & 0 & 1-i \\ 0 & -i & i-\lambda & 1 \\ 1+i & -1+i & 0 & 2i-\lambda \end{vmatrix} = (i-\lambda)\begin{vmatrix} 1-\lambda & i & 1+i \\ -i & 1-\lambda & 1-i \\ 1+i & -1+i & 2i-\lambda \end{vmatrix} =$$

$$= (i-\lambda)[-\lambda^3 + 2(1+i)\lambda^2] = \lambda^2(\lambda - i)(\lambda - 2 - 2i)$$

Como ocurre que este polinomio tiene raíces complejas (no reales), la matriz A no es diagonalizable como raíz real. Para ver si lo es al considerarla como matriz compleja, habrá que comprobar si, para cada autovalor λ_i, coinciden sus multiplicidades algebraica m_i y geométrica d_i. Como esto siempre se verifica para las raíces simples ($\lambda = i$ y $\lambda = 2 + 2i$), sólo habrá que comprobarlo para la raíz doble ($\lambda = 0$). Hallemos los subespacios propios $V_{\lambda=0}$, $V_{\lambda=i}$ y $V_{\lambda=2+2i}$ y comprobemos que la dimensión del primero es 2:

$$V_{\lambda=0}: (A - 0I)X = O \rightarrow \begin{bmatrix} 1 & i & 0 & 1+i \\ -i & 1 & 0 & 1-i \\ 0 & -i & i & 1 \\ 1+i & -1+i & 0 & 2i \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

luego

$$V_{\lambda=0} = \{(x_1, x_2, x_3, x_4) \in \mathbb{C}^4 \,/\, x_1 + ix_2 + (1+i)x_3 = 0,\ -ix_2 + ix_3 + x_4 = 0\}$$
$$= \mathcal{V}((1, i\ 0, -1), (1+i, 0, 1, -i))$$

que tiene dimensión 2.

$$V_{\lambda=i}: (A - iI)X = O: \begin{bmatrix} 1-i & i & 0 & 1+i \\ -i & 1-i & 0 & 1-i \\ 0 & -i & 0 & 1 \\ 1+i & -1+i & 0 & i \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

luego

$$V_{\lambda=i} = \{(x_1, x_2, x_3, x_4) \in \mathbb{C}^4 \,/\, x_1 = 0,\ x_2 = 0,\ x_4 = 0\} = \mathcal{V}(0, 0, 1, 0)$$

$$V_{\lambda=2+2i}: (A - (2+2i)I)X = O: \begin{bmatrix} -1-2i & i & 0 & 1+i \\ -i & -1-2i & 0 & 1-i \\ 0 & -i & -2-i & 1 \\ 1+i & -1+i & 0 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

luego

$$V_{\lambda=2+2i} = \mathcal{V}(5, -5i, 1-2i, 5+5i)$$

Por tanto, A es diagonalizable en el caso complejo y la diagonalización $D = P^{-1}AP$ de A se obtiene para las siguientes matrices D y P:

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & 2+2i \end{bmatrix} \; ; \quad P = \begin{bmatrix} 1 & 1+i & 0 & 5 \\ i & 0 & 0 & -5i \\ 0 & 1 & 1 & 1-2i \\ -1 & -i & 0 & 5+5i \end{bmatrix}$$

9.5. DIAGONALIZACION ORTOGONAL

Según acabamos de decir, la diagonalización de un endomorfismo $f: V \to V$ es un problema que no tiene siempre solución. Nótese también que, cuando f es diagonalizable, no se puede decir nada, con carácter general, acerca de las bases en las que f tiene matriz diagonal, las cuales pueden ser, en principio, cualesquiera.

Pues bien, nos ocuparemos ahora de un caso particular, de especial interés: aquel en el que el endomorfismo f es simétrico. Vamos a comprobar que, entonces, f siempre es diagonalizable y que, además, dicha diagonalización se puede conseguir en una base ortonormal, por lo que se la llama diagonalización ortogonal. Para estos cambios de base, en los que sólo se manejan bases ortonormales, la matriz P de cambio de coordenadas es ortogonal ($P^t = P^{-1}$), por lo que, si sólo se consideran bases ortonormales, decir que dos matrices A y A' son semejantes ($A' = P^{-1}AP$) es equivalente a afirmar que son congruentes ($A' = P^tAP$); por ello, diagonalizar ortogonalmente por semejanza una matriz (simétrica) es lo mismo que diagonalizarla ortogonalmente por congruencia.

ENDOMORFISMOS SIMETRICOS

[175]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial euclídeo $V \neq O$. Se dice que f es un endomorfismo simétrico si $\bar{x} \cdot f(\bar{y}) = f(\bar{x}) \cdot \bar{y}$ para cualesquiera $\bar{x}, \bar{y} \in V$.

Si V tiene dimensión finita y si A es la matriz del endomorfismo f en una base ortonormal de V, entonces f es simétrico si y sólo si A es simétrica.

DEMOSTRACION

Como la base es ortonormal, en ella la matriz métrica del producto escalar es la matriz unidad, $G = I$, con lo que el producto escalar $\bar{x}_1 \cdot \bar{x}_2$, para cualesquiera vectores $\bar{x}_1, \bar{x}_2 \in V$, vendrá dado por:

$$\bar{x}_1 \cdot \bar{x}_2 = X_1^t G X_2 = X_1^t I X_2 = X_1^t X_2$$

donde X_1 y X_2 son las columnas de coordenadas de $\bar{x}_1$ y $\bar{x}_2$. Si las columnas de coordenadas de $\bar{x}, \bar{y} \in V$ son X e Y, las de $f(\bar{x})$ y $f(\bar{y})$ serán AX y AY y, por ello

$$\bar{x} \cdot f(\bar{y}) = X^t AY \quad \text{y} \quad f(\bar{x}) \cdot \bar{y} = (AX)^t Y = X^t A^t Y$$

Resulta entonces que la relación $\bar{x} \cdot f(\bar{y}) = f(\bar{x}) \cdot \bar{y}$ para cualesquiera $\bar{x}, \bar{y} \in V$ equivale a la

$$X^t AY = X^t A^t Y, \quad \forall X, Y \in \mathcal{M}_{n \times 1}$$

Esto último acontece si y sólo si $A = A^t$, esto es, para A simétrica.

EJEMPLO

Sea V el espacio vectorial real de las funciones polinómicas de $[-1, 1]$ en $\mathbb{R}$ cuyo grado es menor o igual que n; en V se considera el producto escalar:

$$(p \mid q) = \int_{-1}^{1} p(x)q(x)\,dx, \quad \forall p.\ q \in V$$

Sea $F: V \to V$ el endomorfismo que a cada función $p \in V$ le asigna la función $F(p) = p^*$ definida por $p^*(x) = p(-x)$ para todo $x \in [-1, 1]$. Este endomorfismo es simétrico ya que:

$$(p \mid F(q)) = \int_{-1}^{1} p(x)q^*(x)\,dx = \int_{-1}^{1} p(x)q(-x)\,dx$$

$$(F(p) \mid q) = \int_{-1}^{1} p^*(t)q(t)\,dt = \int_{-1}^{1} p(-t)q(t)\,dx$$

y estas integrales son iguales (hágase el cambio de variable $x = -t$). Obsérvese que en la base usual de V, esto es, en la $(p_0, p_1, \ldots, p_n)$ donde

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2, \quad \ldots, \quad p_n(x) = x^n$$

la matriz de F es:

$$A = \begin{bmatrix} 1 & & & \\ & -1 & O & \\ & O & \ddots & \\ & & & \pm 1 \end{bmatrix}$$

que es simétrica, a pesar de que dicha base no es ortonormal. La anterior propiedad asegura que, si F es un endomorfismo simétrico, entonces su matriz en cualquier base ortonormal es simétrica; si la base no es ortonormal, la matriz no tiene que ser necesariamente simétrica, pero puede serlo en algunos casos, como el presente.

AUTOVALORES Y AUTOVECTORES DE LOS ENDOMORFISMOS SIMETRICOS

[176]

1. Sea $f: V \to V$ un endomorfismo simétrico, del espacio vectorial euclídeo $V \neq O$. Si λ y μ son dos autovalores de f y si $\bar{x}$ e $\bar{y}$ son vectores propios de f asociados a λ y μ, respectivamente, entonces se verifica que $(\lambda - \mu)\bar{x} \cdot \bar{y} = 0$.
   Sea A una matriz real simétrica. Si λ y μ son raíces (reales o complejas) de la ecuación característica de A y si X e Y son matrices columna tales que $AX = \lambda X$ y $AY = \mu Y$, entonces se verifica que $(\lambda - \mu)X^t Y = O$.
2. Sea $f: V \to V$ un endomorfismo simétrico, en un espacio vectorial euclídeo $V \neq O$ de dimensión finita n. Sea A una matriz real simétrica de tamaño $n \times n$; se puede suponer que A es matriz de f en cierta base ortonormal de V(*). Se verifica que:
   a) Si λ y μ son autovalores distintos de f o de A, entonces los subespacios propios V_λ y V_μ, asociados a λ y μ, son ortogonales(**).
   b) Tanto f como A tienen n autovalores (reales)(***), si se cuenta cada uno tantas veces como indique su multiplicidad algebraica.

(*) Por ejemplo, el endomorfismo de $\mathbb{R}^n$ (euclídeo) asociado a A en la base canónica es simétrico.

(**) Recuérdese que: Para f, es $V_\lambda = \{\bar{x} \in V / f(\bar{x}) = \lambda\bar{x}\}$. Para A, es $V_\lambda = \{X \in \mathcal{M}_{n\times 1} / AX = \lambda X\}$. Dos columnas $X, Y \in \mathcal{M}_{n\times 1}$ son ortogonales si $X^t Y = O$.

(***) Es decir, el polinomio característico $\det(A - \lambda I)$, que siempre tiene n raíces en $\mathbb{C}$ (complejas), aquí es tal que todas ellas son reales.

DEMOSTRACION

1. Se sabe que $f(\bar{x}) = \lambda\bar{x}$ y que $f(\bar{y}) = \lambda\bar{y}$. Multiplicando escalarmente por $\bar{y}$ la primera igualdad y por $\bar{x}$ la segunda y restando, como por ser f simétrico se verifica que $f(\bar{x}) \cdot \bar{y} = \bar{x} \cdot f(\bar{y})$, se obtiene:

$$\left.\begin{aligned} f(\bar{x}) \cdot \bar{y} &= \lambda\bar{x} \cdot \bar{y} \\ \bar{x} \cdot f(\bar{y}) &= \mu\bar{x} \cdot \bar{y} \end{aligned}\right\} \text{ restando: } 0 = (\lambda - \mu)\bar{x} \cdot \bar{y}$$

   Para el caso de la matriz A, hagamos una demostración directa (*). Como $AX = \lambda X$, también se verificará que $X^t A^t = \lambda X^t$; como A es simétrica, será $X^t A = \lambda X^t$. Multiplicando esta igualdad por Y (por la derecha), multiplicando la igualdad $AY = \mu Y$ por X^t (por la izquierda) y restando, se obtiene que:

$$\left.\begin{aligned} X^t AY &= \lambda X^t Y \\ X^t AY &= \mu X^t Y \end{aligned}\right\} \text{ restando: } 0 = (\lambda - \mu)X^t Y$$

(*) No aprovechamos aquí lo que se acaba de obtener para el caso de endomorfismos, pues allí sólo se habló del caso real. Nótese que aquí, para la matriz real simétrica A, los λ y μ pueden ser complejos; también pueden serlo las columnas X e Y.

2. Recurriendo al resultado anterior, se tiene:

a) Para cualesquiera que sean $\bar{x} \in V_\lambda$ e $\bar{y} \in V_\mu$, de acuerdo con el resultado anterior ha de ser $(\lambda - \mu)\bar{x} \cdot \bar{y} = 0$ y, como es $\lambda \neq \mu$, de aquí se deduce que $\bar{x} \cdot \bar{y} = 0$, luego V_λ y V_μ son ortogonales.

b) Supongamos que esta propiedad fuese falsa. esto es, que la ecuación característica $\det(A - \lambda I) = 0$ tuviera una raíz compleja $\lambda_0 = a + bi$ (a y b reales) con $b \neq 0$. De ser así, como los coeficientes del polinomio característico $\det(A - \lambda I)$ son reales, resultaría que el conjugado de λ_0, esto es, $\bar{\lambda}_0 = a - bi$, también sería raíz de la ecuación característica. Como λ_0 es raíz de la ecuación característica, el sistema $AX = \lambda_0 X$ tiene alguna solución no nula $X_0 \in \mathcal{M}_{n\times 1}$, que puede ser compleja, esto es, de la forma $X_0 = X_1 + iX_2$, donde las matrices columna X_1 y X_2 son reales. Nótese que; como $AX_0 = \lambda_0 X_0$, tomando conjugados en esta igualdad, resulta que $A\bar{X}_0 = \bar{\lambda}_0 \bar{X}_0$, donde $\bar{X}_0 = X_1 - iX_2$. Por tanto, recurriendo a la anterior propiedad 1.ª (aquí las anteriores λ y μ son λ_0 y $\bar{\lambda}_0$ y las X e Y son X_0 y $\bar{X}_0$), se obtiene que

$$(\lambda_0 - \bar{\lambda}_0)X_0^t \bar{X}_0 = O, \quad \text{o sea} \quad 2b(X_1^t + iX_2^t)(X_1 - iX_2) = O$$

ahora bien, como $X_1^t X_2$ y $X_2^t X_1$ son el mismo número real, la última igualdad equivale a la:

$$2b(X_1^t X_1 + X_2^t X_2) = O \qquad [1]$$

Obsérvese que $X_1^t X_1$ y $X_2^t X_2$, como son las sumas de los cuadrados de los elementos de X_1 y X_2, respectivamente, son positivos salvo que sea $X_1 = O$ o $X_2 = O$. Como además es $b \neq O$, de la relación [1] se desprende que ha de ser $X_1 = O$ y $X_2 = O$, luego también será $X_0 = X_1 + iX_2 = O$ y esto es falso, pues X_0 era no nulo. Esta contradicción obliga a desechar que λ_0 no sea real.

EJERCICIO

Compruébese con un ejemplo que, dado un endomorfismo $f: V \to V$, en un espacio euclídeo $V \neq O$ de dimensión finita n, para que f sea simétrico no es suficiente con que tenga n autovalores (reales) y que sus subespacios propios sean ortogonales dos a dos.

RESOLUCION

Para verificar la anterior afirmación vale el siguiente ejemplo: tomando $V = \mathbb{R}^3$ con el producto escalar canónico, sea $f: \mathbb{R}^3 \to \mathbb{R}^3$ el endomorfismo

$$(x, y, z) \mapsto (x', y', z') = (3x - 2y, 3y, 5z)$$

Los autovalores de f son $\lambda = 3$ doble y $\lambda = 5$ simple; f tiene, pues, tres autovalores reales. Los subespacios propios de f son

$$V_{\lambda=3} = \{(\alpha, 0, 0) \in \mathbb{R}^3 / \alpha \in \mathbb{R}\} \quad \text{y} \quad V_{\lambda=5} = \{(0, 0, \beta) \in \mathbb{R}^3 / \beta \in \mathbb{R}\}$$

que son ortogonales. Sin embargo, f no es simétrico ya que tomando, por ejemplo, $\vec{u} = (1, 0, 0)$ y $\vec{v} = (0, 1, 0)$, para estos vectores es

$$\vec{u} \cdot f(\vec{v}) \neq \vec{v} \cdot f(\vec{u}), \quad \text{pues } \vec{u} \cdot f(\vec{v}) = -2 \quad \text{y} \quad \vec{v} \cdot f(\vec{u}) = 0$$

TEOREMA ESPECTRAL

[177] Sea $V \neq O$ un espacio vectorial euclídeo de dimensión finita n. Si $f: V \to V$ es un endomorfismo simétrico, entonces existe una base ortonormal de V formada por vectores propios de f.

DEMOSTRACION

Sean $\lambda_1, ..., \lambda_p$ los autovalores (reales) de f; sean m_i y d_i las multiplicidades algebraica y geométrica del autovalor λ_i (para $i = 1, 2, ..., p$). Como f es simétrico, se sabe que $m_1 + \cdots + m_p = n$ y que los subespacios propios V_{λ_i} (para $i = 1, ..., p$) son ortogonales dos a dos. Nótese que f tiene al menos un autovalor (en cuyo caso, su multiplicidad algebraica sería n) y como máximo tiene n (en cuyo caso, serían todos simples).

Tomando una base ortonormal en cada uno de los p subespacios propios V_{λ_i} de f, como éstos son ortogonales entre sí, al unir dichas bases se obtiene un sistema ortonormal de $d_1 + d_2 + \cdots + d_p = d$ vectores, que es base ortonormal de la suma directa

$$U = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_p}$$

Nótese que $d = \dim U$ es el número máximo de autovectores de f linealmente independientes, que se han podido tomar ortonormales. Nótese también que todo vector propio de f, por pertenecer a alguno de los V_{λ_i}, es un vector de U.

Lo que hay que demostrar es que $d = n$ o bien que $U = V$. Suponiendo que esto es falso, es decir, que $d < n$, vamos a obtener una contradicción, lo que probará el teorema:

Sea $U^\perp$ el suplementario ortogonal de U, esto es, el subespacio de V formado por aquellos vectores que son ortogonales a todo vector de U; la dimensión de $U^\perp$ es $d' = n - d > 0$. Veamos que $U^\perp$ es invariante por f, o sea, tal que $f(U^\perp) \subset U^\perp$. Para ello, hay que comprobar que si $\vec{x}' \in U^\perp$ entonces $f(\vec{x}') \in U^\perp$, es decir, que si $\vec{x}' \cdot \vec{x} = 0$ para todo $\vec{x} \in U$ entonces $f(\vec{x}') \cdot \vec{x} = 0$

para todo $\bar{x} \in U$. Así ocurre ya que, si $\bar{x}' \in U^{\perp}$ y $\bar{x} \in U$, como U es la suma de los V_{λ_i}, se puede poner

$$\bar{x} = \bar{x}_1 + \cdots + \bar{x}_p \quad , \quad \text{con } \bar{x}_i \in V_{\lambda_i} \text{ para } i = 1, 2, \ldots, p$$

y como además f es simétrico, se tiene:

$$f(\bar{x}') \cdot \bar{x} = \bar{x}' \cdot f(\bar{x}) = \bar{x}' \cdot f(\bar{x}_1 + \cdots + \bar{x}_p) = \bar{x}' \cdot (f(\bar{x}_1) + \cdots + f(\bar{x}_p)) =$$
$$= \bar{x}' \cdot (\lambda_1 \bar{x}_1 + \cdots + \lambda_p \bar{x}_p) = \lambda_1(\bar{x}' \cdot \bar{x}_1) + \cdots + \lambda_p(\bar{x}' \cdot \bar{x}_p) = \lambda_1 0 + \cdots + \lambda_p 0 = 0$$

puesto que $\bar{x}_1, \ldots, \bar{x}_p$ son ortogonales a $\bar{x}'$, ya que pertenecen a U. Así, pues, todo vector de $U^{\perp}$ se transforma, por f, en otro vector que también es de $U^{\perp}$. Por tanto, la restricción de f a $U^{\perp}$, que denotaremos por la misma letra f, es un endomorfismo de $U^{\perp}$ en $U^{\perp}$. Este endomorfismo $f: U^{\perp} \to U^{\perp}$ es obviamente simétrico, ya que lo es $f: V \to V$. Por tanto, según el teorema anterior, el endomorfismo $f: U^{\perp} \to U^{\perp}$ tiene algún autovector no nulo $\bar{x}_0 \in U^{\perp}$. Esto último es una contradicción, ya que todos los autovectores de $f: U^{\perp} \to U^{\perp}$ lo son de $f: V \to V$ y éstos pertenecen a U, que tiene intersección nula con $U^{\perp}$, luego habría de ser $\bar{x}_0 = \bar{o}$, que es falso.

EJERCICIO

Sea $f: V \to V$ un endomorfismo simétrico, de un espacio vectorial euclídeo $V \neq O$ de dimensión finita. Sean λ_m y λ_M el menor y el mayor de los autovalores de f. Pruébese que, para cualquiera que sea el vector $\bar{u} \in V$, se verifica que:

$$\lambda_m \|\bar{u}\|^2 \leqslant \bar{u} \cdot f(\bar{u}) \leqslant \lambda_M \|\bar{u}\|^2$$

RESOLUCION

Según el teorema espectral, existe una base ortonormal de V formada por vectores propios de f; sea $(\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n)$ dicha base y llamemos $\lambda_1, \lambda_2, \ldots, \lambda_n$ a los autovalores de f asociados, respectivamente, a $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n$ (nótese que los λ_i no tienen que ser, necesariamente, distintos), de manera que $\lambda_m \leqslant \lambda_i \leqslant \lambda_M$ para $i = 1, 2, \ldots, n$. Si $(u_1, u_2, \ldots, u_n)$ son las coordenadas del vector $\bar{u}$ en la base anterior, será:

$$\bar{u} \cdot f(\bar{u}) = \left(\sum_{i=1}^{n} u_i \bar{e}_i\right) \cdot \left(\sum_{j=1}^{n} u_j f(\bar{e}_j)\right) = \left(\sum_{i=1}^{n} u_i \bar{e}_i\right) \cdot \left(\sum_{j=1}^{n} u_j \lambda_j \bar{e}_j\right) =$$

$$= \sum_{i,j=1}^{n} u_i u_j \lambda_j (\bar{e}_i \cdot \bar{e}_j) = \sum_{i=1}^{n} u_i u_i \lambda_i$$

por tanto:

$$\bar{u} \cdot f(\bar{u}) = u_1^2 \lambda_1 + \cdots + u_n^2 \lambda_n \begin{cases} \leqslant (u_1^2 + \cdots + u_n^2)\lambda_M = \|\bar{u}\|^2 \lambda_M \\ \geqslant (u_1^2 + \cdots + u_n^2)\lambda_m = \|\bar{u}\|^2 \lambda_m \end{cases}$$

DIAGONALIZACION ORTOGONAL DE UN ENDOMORFISMO SIMETRICO

[178]

1. Todo endomorfismo simétrico $f: V \to V$, de un espacio vectorial euclídeo $V \neq O$ de dimensión finita, es diagonalizable y lo es ortogonalmente; es decir, existe alguna base ortonormal de V en la que la matriz D de f es diagonal. Una tal base se obtiene reuniendo bases ortonormales de los subespacios propios de f; la diagonal de D está formada por los autovalores de f (cada uno tantas veces como indique su multiplicidad algebraica).

2. Toda matriz real simétrica A, de tamaño $n \times n$, es ortogonalmente diagonalizable, es decir, existe alguna matriz P ortogonal ($P^t = P^{-1}$) tal que $D = P^{-1}AP$ (*). Las columnas de P son vectores propios de A que forman sistema ortonormal; la diagonal de D está formada por los correspondientes autovalores de A (cada uno tantas veces como indique su multiplicidad algebraica).

(*) Como $P^{-1} = P^t$, las matrices A y D son semejantes y, también, congruentes.

COMPROBACION

1. Según afirma el teorema espectral (véase [177]), en V hay una base ortonormal formada por vectores propios de f, lo cual es condición necesaria y suficiente para que f sea diagonalizable (véase [173]) y lo sea en dicha base. Respecto de la matriz D, lo que aquí se dice es ya conocido (véase [174].
2. Para comprobar lo relativo a la matriz A, recurramos al endomorfismo $f: \mathbb{R}^n \to \mathbb{R}^n$ (en $\mathbb{R}^n$ se considera el producto escalar canónico) asociado a la matriz A en la base canónica; como esta base es ortonormal y A es simétrica, f es un endomorfismo simétrico y, por ello, le es de aplicación lo dicho en el punto 1. Así, pues, las columnas de P son las columnas de coordenadas de la base ortonormal en la que f toma forma diagonal y D es la matriz diagonal de los correspondientes autovalores de f, que son los de A.

EJERCICIO

Diagonalizar ortogonalmente el endomorfismo simétrico $f: \mathbb{R}^3 \to \mathbb{R}^3$ dado por

$$f(x, y, z) = (3x + y + z,\ x + 3y + z,\ x + y + 3z)$$

(En $\mathbb{R}^3$ se considera el producto escalar canónico; la base canónica es ortonormal.)

RESOLUCION

El endomorfismo f es, efectivamente, simétrico ya que su matriz A en la base canónica, es simétrica:

$$A=\begin{bmatrix}3 & 1 & 1\\ 1 & 3 & 1\\ 1 & 1 & 3\end{bmatrix}$$

El polinomio característico de A es, según se obtiene fácilmente:

$$\det(A-\lambda I)=-t^3+9t^2-24t+20=-(t-2)^2(t-5)$$

por tanto, los autovalores de f son $\lambda_1=2$ doble y $\lambda_2=5$ simple. Los correspondientes subespacios propios son:

$$V_{\lambda=2}:(A-2I)X=O\rightarrow\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

luego

$$V_{\lambda=2}=\{(x, y, z)\in\mathbb{R}^3 / x+y+z=0\}$$

$$V_{\lambda=5}:(A-5I)X=O\rightarrow\begin{bmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

luego

$$V_{\lambda=5}=\{(\alpha, \alpha, \alpha) / \alpha\in\mathbb{R}\}$$

De entre todas las bases ortonormales de $V_{\lambda=2}$ tomemos una sencilla, como por ejemplo, la $(\bar{e}_1, \bar{e}_2)$, donde

$$\bar{e}_1=\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)\ ;\ \bar{e}_2=\left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$$

una base ortonormal de $V_{\lambda=5}$ es la $(\bar{e}_3)$:

$$\bar{e}_3=\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

Por tanto, en la base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ de $\mathbb{R}^3$ la matriz de f es

$$D=\begin{bmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 5\end{bmatrix}$$

Obsérvese que, al tiempo, se ha obtenido también la diagonalización ortogonal de la matriz simétrica A, que es la $D = P^{-1}AP$, donde P es la matriz ortogonal

$$D = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

EJERCICIO

Sea A una matriz cuadrada que es ortogonalmente diagonalizable, esto es, tal que existen una matriz diagonal D y una matriz ortogonal P tales que $D = P^{-1}AP$. Pruébese que A es simétrica.

RESOLUCION

Se sabe que $P^t = P^{-1}$, pues P es ortogonal. y que $D^t = D$, pues D es diagonal; hay que probar que $A^t = A$. Para ello, expresando que $D^t = D$, se tiene:

$$\left.\begin{aligned} D &= P^{-1}AP = P^tAP \\ D^t &= (P^tAP)^t = P^tA^tP \end{aligned}\right\} \Rightarrow P^tAP = P^tA^tP$$

y multiplicando, en esta última igualdad, por P^{-1} a la derecha y por $(P^t)^{-1} = P$ a la izquierda, se obtiene que $A = A^t$.

OBSERVACIONES

1. Sean A y A' dos matrices reales cuadradas del mismo tamaño. Se dice que A' es *ortogonalmente semejante* a A si A' es semejante a A y lo es con una matriz de cambio ortogonal, esto es, si $A' = P^{-1}AP$ donde P es una matriz regular tal que $P^{-1} = P^t$. Se dice que A' es *ortogonalmente congruente* a A si A' es congruente a A y lo es con una matriz de cambio ortogonal, esto es, si $A' = P^tAP$ donde P es una matriz regular tal que $P^t = P^{-1}$. Es evidente que dos matrices son ortogonalmente semejantes si y sólo si son ortogonalmente congruentes.
2. Según sabemos (véase [178]), toda matriz real simétrica A es ortogonalmente diagonalizable, es decir, se verifica que existe alguna matriz ortogonal P tal que $D = P^{-1}AP = P^tAP$ es una matriz diagonal. Esta matriz D es, al tiempo, diagonalización por semejanza y diagonalización por congruencia de A.

DIAGONALIZACION ORTOGONAL DE UNA FORMA CUADRATICA

[179]

Dada una forma cuadrática real $\omega: V \to \mathbb{R}$, definida en un espacio vectorial euclídeo $V \neq O$ de dimensión finita, existe una base ortonormal de V en la que ω tiene matriz diagonal, D; este resultado se expresa diciendo que ω es ortogonalmente diagonalizable. Si A es la matriz (simétrica) de ω en una base ortonormal de V, dada, entonces: 1.º) la diagonal de D está formada por los autovalores de A (cada uno tantas veces como indique su multiplicidad algebraica); 2.º) es $D = P^t AP$, donde P es matriz ortogonal cuyas columnas son vectores propios de A; 3.º) estas columnas son las de las coordenadas de los vectores de la base de V en la que D es matriz de ω. Si en esta base las coordenadas de un vector $\bar{x}$ se denotan por $(x_1, x_2, ..., x_n)$, será:

$$\omega(\bar{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

($\lambda_1, \lambda_2, ..., \lambda_n$ son los n autovalores (reales) de la matriz A de ω en una base ortonormal de V).

COMPROBACION

Como la matriz A es simétrica, se sabe que (véase [178], 2.º) A es ortogonalmente diagonalizable, esto es, que existe una matriz P ortogonal ($P^t = P^{-1}$) tal que $D = P^t AP$ es diagonal. Como D es congruente con A, resulta que D es matriz de ω en una nueva base. Las columnas de P son las columnas de coordenadas de los vectores de esta base; como la base de partida es ortonormal y la matriz P es ortonormal, resulta que la base donde se produce la diagonalización de ω es también ortonormal. Nótese finalmente que D, por ser la diagonalización ortogonal de A, tiene su diagonal formada por los autovalores de A.

OBSERVACION

Según se acaba de decir, dada una forma cuadrática real $\omega: V \to \mathbb{R}$, donde $V \neq O$ es un espacio vectorial euclídeo de dimensión finita ω, existe una base ortonormal $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V en la que la expresión de $\omega(\bar{x})$ para $\bar{x} \in V$, es:

$$\omega(\bar{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

donde $(x_1, x_2, ..., x_n)$ son las coordenadas de $\bar{x}$ en la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ y $\lambda_1, \lambda_2, ..., \lambda_n$ son los autovalores de la matriz A (simétrica) de ω en cualquier base

ortonormal de V, es decir, los λ_i son las raíces de la ecuación característica $\det(A - \lambda I) = 0$. Así, pues, el rango y la signatura de ω son:

$$\operatorname{rang}\omega = n - m, \quad \text{donde } m = \text{multiplicidad del posible autovalor } \lambda = 0 \text{ de } A \text{ (si } \lambda = 0 \text{ no es autovalor, es } m = 0)$$

$$\operatorname{sig}\omega = (p, q), \quad \text{donde } \begin{cases} p = \text{número de autovalores positivos de } A \\ q = \text{número de autovalores negtivos de } A \end{cases}$$

(cada autovalor se cuenta tantas veces como indica su multiplicidad).

Nótese que, para esto de la signatura, no se necesita conocer los autovalores de A, sino que basta con saber cuántos hay positivos (los negativos q son $q = n - p - m$). Para esto último es de interés el teorema de Descartes, sobre el número de raíces positivas que puede tener una ecuación algebraica. En nuestro caso, llamando $a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n$ al polinomio característico, como éste tiene sus n raíces reales (pues A es simétrica), el teorema de Descartes permite asegurar que el número de autovalores positivos de A es igual al número total de variaciones de signo que se presentan entre cada dos elementos consecutivos de la sucesión $a_0, a_1, a_2, \ldots, a_n$, de los coeficientes del polinomio característico de A; si alguno de estos coeficientes a_i fuese nulo, se le puede suponer positivo o negativo (debido a que a_{i-1} y a_{i+1} tienen aquí distinto signo).

γ DIAGONALIZACION POR BLOQUES DE UNA TRANSFORMACION ORTOGONAL

Vamos a ocuparnos aquí de una transformación ortogonal $f: V \to V$, donde $V \neq O$ es un espacio vectorial euclídeo de dimensión finita n. Aunque, en general, la transformación ortogonal f no es diagonalizable, es, sin embargo, «diagonalizable por bloques» de tamaños 1×1 y 2×2 y lo es ortogonalmente. Es decir, existe alguna base ortonormal de V en la que f tiene matriz de la forma

$$\begin{bmatrix} A_1 & & & \\ & A_2 & & O \\ & & \ddots & \\ & O & & A_p \end{bmatrix}$$

donde las submatrices A_i, que ocupan posiciones diagonales en A, tienen tamaños 1×1 ó 2×2; más exactamente, son de los tipos:

$$A_i = [1] \qquad A_i = \begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$$
$$A_i = [-1]$$

Para llegar a este resultado, se va a seguir un camino directo, conciso y que precisa de pocos conocimientos adicionales.

9.6. SUBESPACIOS INVARIANTES DE UNA TRANSFORMACION ORTOGONAL

AUTOVALORES DE UNA TRANSFORMACION ORTOGONAL

[180]

Si $f: V \to V$ es una transformación ortogonal, del espacio vectorial euclídeo $V \neq O$ de dimensión finita n, se verifica que:

1. Si f tiene algún autovalor λ_0 (real), éste ha de ser $\lambda_0 = 1$ ó $\lambda_0 = -1$. Si n es impar, entonces f admite como autovalor a uno, al menos, de los $\lambda_0 = 1$ ó $\lambda_0 = -1$.
2. Si $\lambda_0 \in \mathbb{C}$ es un autovalor complejo(*) de f, entonces es $|\lambda_0| = 1$, o sea, $\lambda_0 = \cos\alpha + i \operatorname{sen}\alpha$(**) para cierto $\alpha \in]-\pi, \pi]$ (para $\alpha = 0$, es $\lambda_0 = 1$; para $\alpha = \pi$, es $\lambda_0 = -1$; para los demás valores de α, λ_0 no es real).

(*) O sea, λ_0 es una raíz compleja de la ecuación característica de f.

(**) Nótese que $\bar{\lambda}_0 = \cos\alpha - i \operatorname{sen}\alpha$ también es raíz de la ecuación característica pues ésta tiene sus coeficientes reales.

DEMOSTRACION

Aunque el apartado 1 puede considerarse como un caso particular del 2, vamos a demostrarlo separadamente para independizarlo del caso complejo.

1. Como $\lambda_0 \in \mathbb{R}$ es un autovalor de f, existe un vector no nulo $\bar{x} \in V$ tal que $f(\bar{x}) = \lambda_0 \bar{x}$. Ahora bien, por ser f ortogonal, se verifica que $\|f(\bar{x})\| = \|\bar{x}\|$; por tanto, $\|\lambda_0 \bar{x}\| = \|\bar{x}\|$, o sea, $|\lambda_0| \, \|\bar{x}\| = \|\bar{x}\|$ y, como $\|\bar{x}\| \neq 0$ por ser $\bar{x} \neq \bar{o}$, resulta que $|\lambda_0| = 1$, luego ha de ser $\lambda_0 = 1$ o $\lambda_0 = -1$.
   Si n es impar, entonces la ecuación característica de f tiene al menos una raíz real λ_0, ya que dicha ecuación tiene coeficientes reales y grado impar. Este valor $\lambda_0 \in \mathbb{R}$ es, pues, un autovalor (real) de f y, según se acaba de probar, debe entonces ser $\lambda_0 = 1$ o $\lambda_0 = -1$.
2. Recurramos a una base ortonormal de V y sea A la matriz de f en dicha base; esta matriz es ortogonal, esto es, tal que $A^t = A^{-1}$. Sea $\lambda_0 \in \mathbb{C}$ una raíz del polinomio característico de f, es decir, de $\det(A - \lambda I)$. Se sabe entonces que el sistema de ecuaciones $(A - \lambda_0 I)X = O$ tiene alguna solución no nula $X_0 \neq O$, que tendrá, en general, sus elementos complejos. Por verificarse que $(A - \lambda_0 I)X_0 = O$, o sea, que $AX_0 = \lambda X_0$, también se cumple que $A\bar{X}_0 = \bar{\lambda}_0 \bar{X}_0$, donde $\bar{\lambda}_0$ es el complejo conjugado de λ_0 y $\bar{X}_0$ es la matriz

columna cuyos elementos son los conjugados de los respectivos elementos de X_0, es decir:

$$\text{si} \quad X_0 = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}, \quad \text{entonces } \bar{X}_0 = \begin{bmatrix} \bar{z}_1 \\ \bar{z}_2 \\ \vdots \\ \bar{z}_n \end{bmatrix}$$

A partir de $AX_0 = \lambda_0 X_0$, trasponiendo y recurriendo a que $A^t = A^{-1}$, se obtiene:

$$AX_0 = \lambda_0 X_0 \Rightarrow X_0^t A^t = \lambda_0 X_0^t \Rightarrow X_0^t A^{-1} = \lambda_0 X_0^t$$

Multiplicando ahora miembro a miembro esta última igualdad y la $A\bar{X}_0 = \bar{\lambda}_0 \bar{X}_0$, se obtiene:

$$X_0^t A^{-1} A \bar{X}_0 = \lambda_0 \bar{\lambda}_0 X_0^t \bar{X}_0, \quad \text{o sea} \quad X_0^t \bar{X}_0 = |\lambda_0|^2 X_0^t \bar{X}_0 \qquad [1]$$

Ahora bien, como $X_0 \neq O$, resulta que

$$X_0^t \bar{X}_0 = z_1 \bar{z}_1 + \cdots + z_n \bar{z}_n = |z_1|^2 + \cdots + |z_n|^2 > 0$$

luego $X_0^t \bar{X}_0 \neq 0$ y de la relación [1] se deduce que $|\lambda_0|^2 = 1$, es decir, la raíz $\lambda_0 \in \mathbb{C}$ del polinomio característico tiene módulo unidad.

EJEMPLO

Considérese la siguiente matriz A, que es ortogonal:

$$A = \begin{bmatrix} 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \\ -1/3 & 2/3 & 2/3 \end{bmatrix}$$

El polinomio característico de A será el:

$$\begin{vmatrix} 2/3-\lambda & -1/3 & 2/3 \\ 2/3 & 2/3-\lambda & -1/3 \\ -1/3 & 2/3 & 2/3-\lambda \end{vmatrix} = -\lambda^3 + 2\lambda^2 - 2\lambda + 1 = -(\lambda - 1)(\lambda^2 - \lambda + 1)$$

Las raíces de este polinomio son:

$$\lambda_1 = 1, \quad \lambda_2 = \frac{1}{2} + i\frac{\sqrt{3}}{2} = \cos 60^\circ + i \operatorname{sen} 60^\circ$$

$$\lambda_3 = \bar{\lambda}_2 = \frac{1}{2} - i\frac{\sqrt{3}}{2} = \cos 60^\circ - i \operatorname{sen} 60^\circ = \cos(-60^\circ) + i \operatorname{sen}(-60^\circ)$$

que responden a lo previsto en la anterior proposición.

SUBESPACIOS BIDIMENSIONALES INVARIANTES

[181]

Sea $f: V \to V$ una transformación ortogonal, del espacio vectorial euclídeo $V \neq O$ de dimensión finita. Se verifica que:

Si $\lambda_0 = \cos\alpha + i \operatorname{sen}\alpha$ es una raíz compleja (no real) de la ecuación característica de f, entonces existen dos vectores $\bar{x}_0$ e $\bar{y}_0$ de V que forman sistema ortonormal y son tales que

$$\left.\begin{aligned} f(\bar{x}_0) &= (\cos\alpha)\bar{x}_0 - (\operatorname{sen}\alpha)\bar{y}_0 \\ f(\bar{y}_0) &= (\operatorname{sen}\alpha)\bar{x}_0 + (\cos\alpha)\bar{y}_0 \end{aligned}\right\} (*)$$

El subespacio $V_0 = \mathcal{V}(\bar{x}_0, \bar{y}_0)$, que engendran $\bar{x}_0$ e $\bar{y}_0$, es invariante por f; se verifica que $f(V_0) = V_0$. La restricción de f a V_0 es una rotación de ángulo α.

(*) Si se toma una base ortonormal en V y si A es la matriz de f en dicha base, $\bar{x}_0$ e $\bar{y}_0$ tienen por columnas de coordenadas a X_0 e Y_0 soluciones de

$$[A - (\cos\alpha + i \operatorname{sen}\alpha)I](X_0 + iY_0) = O$$

DEMOSTRACION

1. Sea A la matriz (ortogonal) de f en una cierta base ortonormal de V. Como λ_0 es raíz de $\det(A - \lambda I) = 0$, ecuación característica de f, resulta que la ecuación matricial $(A - \lambda_0 I)Z = O$ tiene alguna solución no nula Z_0 (matriz columna) que puede tener sus elementos complejos, por lo que se la expresará en la forma $Z_0 = X_0 + iY_0$, donde $X_0, Y_0 \in \mathcal{M}_{n\times 1}$ son reales. Así, pues, existen dos matrices columna reales tales que:

$$A(X_0 + iY_0) = \lambda_0(X_0 + iY_0) = (\cos\alpha + i \operatorname{sen}\alpha)(X_0 + iY_0) =$$

$$= (\cos\alpha X_0 - \operatorname{sen}\alpha Y_0) + i(\operatorname{sen}\alpha X_0 + \cos\alpha Y_0)$$

lo que equivale a las dos igualdades (reales):

$$AX_0 = \cos\alpha X_0 - \operatorname{sen}\alpha Y_0 \quad \text{y} \quad AY_0 = \operatorname{sen}\alpha X_0 + \cos\alpha Y_0$$

Si $\bar{x}_0, \bar{y}_0 \in V$ son los vectores que tienen por columnas de coordenadas a las X_0 e Y_0, las anteriores igualdades se pueden expresar poniendo:

$$f(\bar{x}_0) = (\cos\alpha)\bar{x}_0 - (\operatorname{sen}\alpha)\bar{y}_0 \quad \text{y} \quad f(\bar{y}_0) = (\operatorname{sen}\alpha)\bar{x}_0 + (\cos\alpha)\bar{y}_0 \qquad [1]$$

2. Veamos ahora que los vectores $\bar{x}_0$ e $\bar{y}_0$ son ortogonales, tienen la misma norma y pueden conseguirse unitarios, con lo que $(\bar{x}_0, \bar{y}_0)$ será un sistema ortonormal. Como f es ortogonal, de las relaciones [1] se deduce que:

$$\|\bar{x}_0\|^2 = \|f(\bar{x}_0)\|^2 = \cos^2\alpha\|\bar{x}_0\|^2 + \operatorname{sen}^2\alpha\|\bar{y}_0\|^2 - 2\operatorname{sen}\alpha\cos\alpha(\bar{x}_0 \cdot \bar{y}_0)$$
$$\bar{x}_0 \cdot \bar{y}_0 = f(\bar{x}_0) \cdot f(\bar{y}_0) = \operatorname{sen}\alpha\cos\alpha(\|\bar{x}_0\|^2 - \|\bar{y}_0\|^2) + (\cos^2\alpha - \operatorname{sen}^2\alpha)\bar{x}_0 \cdot \bar{y}_0$$

y estas igualdades son respectivamente equivalentes a las:

$$(\|\bar{y}_0\|^2 - \|\bar{x}_0\|^2)\operatorname{sen}\alpha = 2(\bar{x}_0 \cdot \bar{y}_0)\cos\alpha \qquad [2]$$
$$2(\bar{x}_0 \cdot \bar{y}_0)\operatorname{sen}\alpha = -\cos\alpha(\|\bar{y}_0\|^2 - \|\bar{x}_0\|^2) \qquad [3]$$

Multiplicando las igualdades [2] y [3] y pasando, luego, todo al primer miembro, resulta que:

$$2(\bar{x}_0 \cdot \bar{y}_0)(\|\bar{y}_0\|^2 - \|\bar{x}_0\|^2) = 0$$

y, por tanto, ha de ser $\bar{x}_0 \cdot \bar{y}_0 = 0$ ó $\|\bar{y}_0\| = \|\bar{x}_0\|$. Si ocurre lo primero, de [2] se desprende que también ocurre lo segundo. Si se verifica esto último, de [3] se deduce que entonces es $\bar{x}_0 \cdot \bar{y}_0 = 0$. Por tanto, es:

$$\bar{x}_0 \cdot \bar{y}_0 = 0 \quad \text{y} \quad \|\bar{x}_0\| = \|\bar{y}_0\|$$

Obsérvese que $\bar{x}_0$ e $\bar{y}_0$ se pueden suponer unitarios, con lo que $(\bar{x}_0, \bar{y}_0)$ sería un sistema ortonormal, pues si no lo fuesen, sustituyendo $\bar{x}_0$ e $\bar{y}_0$ por $\bar{x}_0/\|\bar{x}_0\|$ e $\bar{y}_0/\|y_0\|$, que forman sistema ortonormal, se seguirían verificando las relaciones [1].

3. El subespacio $V_0 = \mathcal{V}(\bar{x}_0, \bar{y}_0)$ es invariante por f ya que, como $\bar{x}'_0 = f(\bar{x}_0)$ e $\bar{y}'_0 = f(\bar{y}_0)$ han resultado ser combinaciones lineales de $\bar{x}_0$ e $\bar{y}_0$, esto es $\bar{x}'_0$, $\bar{y}'_0 \in V_0$, nos encontramos con que todo vector $\bar{v} \in V_0$, esto es, cualquier $\bar{v} = h\bar{x}_0 + k\bar{y}_0$ (con $h, k \in \mathbb{R}$) se transforma en un vector que también es de V_0:

$$f(\bar{v}) = f(h\bar{x}_0 + k\bar{y}_0) = h\bar{x}'_0 + k\bar{y}'_0 \in V_0$$

Como f es biyectiva (por ser ortogonal) y V_0 tiene dimensión 2, también $f(V_0)$ tiene dimensión 2 y, por ello, al ser $f(V_0) \subset V_0$, ha de ser $f(V_0) = V_0$.

4. La aplicación $f: V_0 \to V_0$ (nótese que V_0 es invariante por f) es una rotación de ángulo α (véase [152]) pues, en la base ortonormal $(\bar{x}_0, \bar{y}_0)$ de V_0, su matriz es, de acuerdo con las relaciones [1]:

$$\begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$$

EJERCICIO

Sea $f: V \to V$ una transformación ortogonal, del espacio vectorial euclídeo $V \neq O$ de dimensión finita. Se sabe que $\lambda_0 = \cos\alpha + i\operatorname{sen}\alpha$ (con $\operatorname{sen}\alpha \neq 0$) y

$\lambda_1 = 1$ son dos raíces del polinomio característico de f. Sea V_0 el subespacio de V engendrado por los vectores ortonormales $\bar{x}_0, \bar{y}_0 \in V$ tales que

$$f(\bar{x}_0) = (\cos\alpha)\bar{x}_0 - (\operatorname{sen}\alpha)\bar{y}_0 \quad \text{y} \quad f(\bar{y}_0) = (\operatorname{sen}\alpha)\bar{x}_0 + (\cos\alpha)\bar{y}_0 \qquad [1]$$

(que existen según la proposición precedente). Sea V_1 el subespacio propio de f asociado a su autovalor $\lambda_1 = 1$. Pruébese que los subespacios V_0 y V_1 son ortogonales entre sí.

RESOLUCION

Como V_0 tiene dimensión 2 y $(\bar{x}_0, \bar{y}_0)$ es una de sus bases, habrá que comprobar que todo vector $\bar{u}_1 \in V_1$ es ortogonal a $\bar{x}_0$ y a $\bar{y}_0$, esto es, que $\bar{u}_1 \cdot \bar{x}_0 = 0$ y $\bar{u}_1 \cdot \bar{y}_0 = 0$. Para ello, recurriendo a las relaciones [1], teniendo en cuenta que $f(\bar{u}_1) = \bar{u}_1$ y por ser f ortogonal (conserva el producto escalar), podemos poner:

$$\bar{u}_1 \cdot \bar{x}_0 = f(\bar{u}_1) \cdot f(\bar{x}_0) = \bar{u}_1 \cdot (\bar{x}_0 \cos\alpha - \bar{y}_0 \operatorname{sen}\alpha) = (\bar{u}_1 \cdot \bar{x}_0)\cos\alpha - (\bar{u}_1 \cdot \bar{y}_0)\operatorname{sen}\alpha$$
$$\bar{u}_1 \cdot \bar{y}_0 = f(\bar{u}_1) \cdot f(\bar{y}_0) = \bar{u}_1 \cdot (\bar{x}_0 \operatorname{sen}\alpha + \bar{y}_0 \cos\alpha) = (\bar{u}_1 \cdot \bar{x}_0)\operatorname{sen}\alpha + (\bar{u}_1 \cdot \bar{y}_0)\cos\alpha$$

Este sistema homogéneo de dos ecuaciones, en las dos incógnitas $\bar{u}_1 \cdot \bar{x}_0$ y $\bar{u}_1 \cdot \bar{y}_0$, tiene por determinante a

$$\begin{vmatrix} \cos\alpha - 1 & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha - 1 \end{vmatrix} = (\cos\alpha - 1)^2 + \operatorname{sen}^2\alpha = 2(\cos\alpha + 1)$$

que no puede ser nulo, ya que $\operatorname{sen}\alpha \neq 0$ y entonces $\cos\alpha \neq -1$. Como el determinante es no nulo, el sistema homogéneo sólo tiene la solución nula, luego ha de ser $\bar{u}_1 \cdot \bar{x}_0 = 0$ y $\bar{u}_1 \cdot \bar{y}_0 = 0$.

PROPOSICION

[182]

Sea $f: V \to V$ una transformación ortogonal, del espacio vectorial euclídeo $V \neq O$ de dimensión finita. Si U es un subespacio vectorial de V que es invariante por f, entonces el subespacio $U^\perp$, suplementario ortogonal de U, también es invariante por f.

DEMOSTRACION

Hay que probar que si $\bar{v} \in U^\perp$ entonces $f(\bar{v}) \in U^\perp$; es decir que, dado $\bar{v} \in V$ tal que $\bar{v} \cdot \bar{u} = 0$ para todo $\bar{u} \in U$, entonces $f(\bar{v}) \cdot \bar{u} = 0$ para todo $\bar{u} \in U$. Para cualquiera que sea $\bar{u} \in U$, como f es biyectiva y U es invariante por f, existe un $\bar{u}' \in U$ tal que $f(\bar{u}') = \bar{u}$. Entonces, como f es ortogonal, para todo $\bar{u} \in U$, será:

$$f(\bar{v}) \cdot \bar{u} = f(\bar{v}) \cdot f(\bar{u}') = \bar{v} \cdot \bar{u}' = 0,$$

como había que comprobar.

9.7. FORMA CANONICA DE UNA TRANSFORMACION ORTOGONAL

PROPOSICION

[183] Dada una transformación ortogonal $f: V \to V$, del espacio vectorial euclídeo $V \neq O$ de dimensión finita, existe alguna base ortonormal de V en la que f tiene matriz de la forma:

$$C = \begin{pmatrix} \begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix} & & & & O \\ & \begin{matrix} -1 & & O \\ & \ddots & \\ O & & -1 \end{matrix} & & & \\ & & \begin{matrix} \cos\alpha_1 & -\operatorname{sen}\alpha_1 \\ \operatorname{sen}\alpha_1 & \operatorname{sen}\alpha_1 \end{matrix} & & \\ & & & \ddots & \\ & & & & \begin{matrix} \cos\alpha_p & -\operatorname{sen}\alpha_p \\ \operatorname{sen}\alpha_p & \cos\alpha_p \end{matrix} \\ O & & & & \end{pmatrix}$$

Esta matriz C diagonal por bloques es única, salvo el orden en el que figuran en ella los bloques diagonales, pero en general no es única la base en la que C es matriz de f.

NOTA: el primero de los bloques es una matriz unidad; el segundo es la matriz opuesta de una matriz unidad; los siguientes bloques son todos de tamaño 2×2.

DEMOSTRACION

Procedamos por inducción sobre $n = \dim V$. Para $n = 1$, la propiedad se verifica obviamente, pues en cualquier base ortonormal la matriz de f es forzosamente la [1] o la [−1]. Para $n = 2$ la propiedad también es cierta, como se probó anteriormente (véase [152]). Supongamos que la propiedad se verifica para toda transformación ortogonal de un espacio vectorial euclídeo de dimensión menor que n. Hemos de comprobar que, entonces, la propiedad también se verifica si la dimensión del espacio es n (con $n \geq 3$).

La ecuación característica de f tiene al menos una raíz real o compleja (*). Esta raíz es (véase [180]):

$$\lambda_0 = 1 \quad \text{ó} \quad \lambda_0 = -1 \quad \text{ó} \quad \lambda_0 = \cos\alpha + i\,\text{sen}\,\alpha \quad (\text{con } \text{sen}\,\alpha \neq 0)$$

analicemos cada una de estas posibilidades:

1. Si $\lambda_0 = \pm 1$, entonces existe algún vector no nulo $\bar{u}_0 \in V$, que se puede suponer unitario, tal que $f(\bar{u}_0) = \pm\bar{u}_0$ y, por tanto, el subespacio $U = \mathcal{V}(\bar{u}_0)$, engendrado por $\bar{u}_0$, es invariante. Según ya sabemos (véase [182]), también es entonces invariante el subespacio $U^\perp$, suplementario ortogonal de U. Como $\dim U^\perp = n - 1 \geqslant 2$, y según la hipótesis de inducción aplicada a la transformación ortogonal $f: U^\perp \to U^\perp$, existe una base ortonormal $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_{n-1})$ de $U^\perp$ en la que la matriz C_{n-1} de esta transformación es del tipo descrito en el enunciado. Nótese que $(\bar{u}_0, \bar{u}_1, \bar{u}_2, \ldots, \bar{u}_{n-1})$ es una base ortonormal de V y que, en ella, la matriz de $f: V \to V$ es la C_n siguiente:

$$C_n = \left[\begin{array}{c|c} \pm 1 & O \\ \hline O & C_{n-1} \end{array}\right]$$

Esta matriz, salvo quizá el orden de colocación de sus elementos y bloques diagonales, es del tipo descrito en el enunciado. Por tanto, sin más que permutar el orden de los vectores de la anterior base, si es que es preciso, se obtiene una matriz C del tipo buscado.

2. Si $\lambda_0 = \cos\alpha + i\,\text{sen}\,\alpha$ (con $\text{sen}\,\alpha \neq 0$), entonces (véase [181]) existe un sistema ortonormal $(\bar{x}_0, \bar{y}_0)$ de vectores de V tal que el subespacio $U = \mathcal{V}(\bar{x}_0, \bar{y}_0)$ es invariante por f y, además, la matriz asociada a la transformación $f: U \to U$ en la base ortonormal $(\bar{x}_0, \bar{y}_0)$ es la:

$$\begin{bmatrix} \cos\alpha & -\text{sen}\,\alpha \\ \text{sen}\,\alpha & \cos\alpha \end{bmatrix}$$

Según se probó anteriormente (véase [182]), por ser U invariante por f, también lo es $U^\perp$, suplementario ortogonal de U. La dimensión de $U^\perp$ es $n - 2 \geqslant 1$ y, por ello, a la transformación ortogonal $f: U^\perp \to U^\perp$ le es aplicable la hipótesis de inducción, de la que se deduce que existe alguna base ortonormal $(\bar{u}_1, \ldots, \bar{u}_{n-2})$ de $U^\perp$ en la que la matriz C_{n-2} de esta transformación ortogonal es de la forma que dice el enunciado. Recurriendo a que $(\bar{x}_0, \bar{y}_0, \bar{u}_1, \ldots, \bar{u}_{n-2})$ es una base ortonormal de V y a que, en ella, la matriz de $f: V \to V$ es entonces la:

(*) La ecuación característica tiene n raíces en C, pero pudieran ser todas iguales.

$$C_n = \left[\begin{array}{cc|c} \cos\alpha & -\operatorname{sen}\alpha & \\ \operatorname{sen}\alpha & \cos\alpha & O \\ \hline & O & C_{n-1} \end{array}\right]$$

nos encontramos con que esta matriz C_n de f es, salvo quizá el orden de colocación de sus elementos y bloques diagonales, del tipo buscado. Resulta pues que, con sólo cambiar el orden de los vectores de la anterior base, si fuera necesario, se obtiene que la matriz de f en cierta base ortonormal de V es una matriz C de las del enunciado.

OBSERVACION

La anterior forma canónica C (véase [183]), de la transformación ortogonal $f: V \to V$ (donde $V \neq O$ es euclídeo y tiene dimensión finita n), consta de los siguientes bloques diagonales, cuyos tamaños son los indicados, como subíndices:

$$I_{h\times h},\quad -I_{k\times k},\quad [\alpha_i]_{2\times 2} = \begin{bmatrix} \cos\alpha_i & -\operatorname{sen}\alpha_i \\ \operatorname{sen}\alpha_i & \cos\alpha_i \end{bmatrix} \quad \text{para } i = 1, 2, \ldots, p$$

(se verifica que $h + k + 2p = n$).

Nótese que f y C tienen el mismo polinomio característico, y, por ello, las n raíces (reales y complejas) del polinomio característico de C, que son

$\lambda = 1$ (con multiplicidad algebraica h)

$\lambda = -1$ (con multiplicidad algebraica k)

$\lambda = \cos\alpha_i \pm i \operatorname{sen}\alpha_i$ (para $i = 1, 2, \ldots, p$)

son también las del polinomio característico de f, que sólo dependen de f. Por tanto, h, k, α_1, α_2, ..., α_p son únicos, lo que permite asegurar que la matriz canónica C, de la transformación ortogonal f, es única, salvo que se altere el orden en el que se ponen los bloques $[\alpha_1]$. $[\alpha_2]$, ..., $[\alpha_p]$.

❑ PROPOSICION

El proceso que nos ha conducido a la forma canónica C de una transformación ortogonal f no proporciona un método práctico para determinar, además de la matriz C, una base en la que C sea matriz de f. Para resolver esta cuestión, vamos a precisar de los siguientes resultados:

[184]

Dada una transformación ortogonal $f: V \to V$, del espacio vectorial euclídeo $V \neq O$ de dimensión finita n, y llamando $g: V \to V$ al endomorfismo $g = (1/2)(f + f^{-1})$ (*), se verifica que: si la forma canónica ortogonal de f es la siguiente matriz C, entonces la diagonalización ortogonal del endomorfismo simétrico g se produce en la misma base ortogonal en la que C es matriz de f y, en dicha base, la matriz de g es la siguiente matriz D:

$$C = \begin{bmatrix} I_{h\times h} & & & & O \\ & -I_{k\times k} & & & \\ & & [\alpha_1] & & \\ & O & & \ddots & \\ & & & & [\alpha_p] \end{bmatrix}$$

$$D = \begin{bmatrix} I_{h\times h} & & & & & O \\ & -I_{k\times k} & & & & \\ & & \cos\alpha_1 & & & \\ & & & \cos\alpha_1 & & \\ & & & & \ddots & \\ & O & & & & \cos\alpha_p \\ & & & & & \cos\alpha_p \end{bmatrix}$$

$$\text{con} \quad [\alpha_i] = \begin{bmatrix} \cos\alpha_i & -\operatorname{sen}\alpha_i \\ \operatorname{sen}\alpha_i & \cos\alpha_i \end{bmatrix} \quad \operatorname{sen}\alpha_i \neq 0, \quad h + k + 2p = n$$

Además, si $\bar{u} \in V$ es un vector no nulo del subespacio propio de g correspondiente a uno de sus autovalores $\lambda = \cos\alpha_i$, entonces el subespacio bidimensional $\mathcal{V}(\bar{u}, f(\bar{u}))$ es invariante por f.

(*) El endomorfismo g es simétrico ya que, si es A la matriz ortogonal de f en una base ortonormal de V, entonces la matriz de g en dicha base es la $(1/2)(A + A^{-1}) = (1/2)(A + A^t)$, que es simétrica.

DEMOSTRACION

Según se ha indicado en la nota, recurriendo a una base ortonormal de V y si A es la matriz ortogonal ($A^t = A^{-1}$) de f en dicha base, la matriz de g es:

$$B = \tfrac{1}{2}(A + A^{-1}) = \tfrac{1}{2}(A + A^t)$$

que es simétrica.

Se sabe que existe una matriz ortogonal P tal que $C = P^{-1}AP$ y, por tanto,

$$P^{-1}BP = \tfrac{1}{2}\,P^{-1}(A + A^t)P = \tfrac{1}{2}\,P^{-1}AP + \tfrac{1}{2}\,P^{-1}A^tP = \tfrac{1}{2}\,P^{-1}AP + \tfrac{1}{2}\,(P^tAP^{-1t})^t =$$

$$= \tfrac{1}{2}\,C + \tfrac{1}{2}\,(P^{-1}AP)^t = \tfrac{1}{2}\,C + \tfrac{1}{2}\,C^t = D$$

Por tanto, con la misma matriz de cambio ortogonal P que transforma (por semejanza) A en C, la matriz B se transforma en D.

Para demostrar la última afirmación del enunciado, bastará con comprobar que f transforma a $\bar{u}$ y a $f(\bar{u})$ en combinaciones lineales de $\bar{u}$ y $f(\bar{u})$, esto es, que $f(f(\bar{u}))$ depende linealmente de $\bar{u}$ y $f(\bar{u})$. Así ocurre, ya que como se sabe que $g(\bar{u}) = \cos\alpha_i\bar{u}$, será $(f + f^{-1})(\bar{u}) = 2\cos\alpha_i\bar{u}$, y hallando la imagen por f de ambos miembros de esta igualdad, será:

$$f(f(\bar{u}) + f^{-1}(\bar{u})) = f(2\cos\alpha_i\bar{u}), \quad \text{o sea} \quad f(f(\bar{u})) + \bar{u} = 2\cos\alpha_i f(\bar{u})$$

como se quería comprobar.

OBTENCION DE LA FORMA CANONICA DE UNA TRANSFORMACION ORTOGONAL

A la vista de los resultados obtenidos anteriormente, en [183] y [184], resulta ya evidente que para diagonalizar ortogonalmente por bloques una transformación ortogonal, se puede proceder del modo que aquí se describe:

[185] Para obtener la forma canónica de una transformación ortogonal $f: V \to V$, del espacio vectorial euclídeo $V \neq O$ de dimensión finita n, conocida su matriz ortogonal A en una base ortonormal de V, recurriremos al endomorfismo simétrico g cuya matriz, en dicha base, es $B = (1/2)(A + A^t)$. Empezaremos por diagonalizar ortogonalmente el endomorfismo g, obteniendo su matriz diagonal D y la correspondiente base $\mathcal{B}$. La matriz D constará de las siguientes submatrices diagonales: 1) una matriz unidad I. de tamaño $h \times h$ (con $h \geqslant 0$); 2) una matriz $-I$, de tamaño $k \times k$ (con $k \geqslant 0$); 3) cierto número de matrices del tipo $(\cos\alpha)I$, de tamaño par $(2r \times 2r)$.

Los bloques diagonales que constituyen la forma canónica C y la correspondiente base de V son:

- En primer y segundo lugar, en C figuran los bloques $I_{h\times h}$ y $-I_{k\times k}$; para conseguirlo, los $h + k$ primeros vectores de la base serán los mismos que en $\mathcal{B}$, es decir, la unión de una base ortonormal de cada uno de los subespacios propios $V_{\lambda=1}$ y $V_{\lambda=-1}$ de g.

- Después, por cada autovalor $\lambda = \cos\alpha$ de g, que tiene multiplicidad par $2m$, en C figura el siguiente bloque diagonal de tamaño $2m \times 2m$:

$$\begin{bmatrix} [\alpha] & & O \\ & [\alpha] & \\ O & & \ddots \\ & & & [\alpha] \end{bmatrix} \quad \text{donde } [\alpha] = \begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$$

Para conseguir este bloque, la base debe tener los correspondientes $2m$ vectores formando una base ortonormal del subespacio propio $V_{\cos\alpha}$ de g, pero esta base no puede ser cualquiera. Para obtener una base válida, se toma en $V_{\cos\alpha}$ un vector $\bar{u} \neq \bar{o}$ cualquiera, se halla $f(\bar{u})$ y se selecciona una base ortonormal cualquiera $(\bar{e}_1, \bar{e}_2)$ del subespacio $\mathcal{V}(\bar{u}, f(\bar{u}))$; después se toma en $V_{\cos\alpha}$ un segundo vector $\bar{v} \neq \bar{o}$ ortogonal a los $\bar{e}_1$ y $\bar{e}_2$, se halla $f(\bar{v})$ y se elige una base ortonormal cualquiera $(\bar{e}_3, \bar{e}_4)$ del subespacio $\mathcal{V}(\bar{v}, f(\bar{v}))$; reiterando el proceso m veces se obtiene una base ortonormal $(\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_{2m})$ en la que C es la matriz de f.

EJERCICIO

Obtener la forma canónica de la transformación ortogonal $f: \mathbb{R}^5 \to \mathbb{R}^5$, en la que la imagen del vector $(x_1, x_2, x_3, x_4, x_5)$ es el vector $(x_1', x_2', x_3', x_4', x_5')$, siendo:

$$x_1' = \frac{3x_1 - 2x_2 - 2x_4 - x_5}{3\sqrt{2}}, \quad x_2' = \frac{2x_1 + 3x_2 - x_4 + 2x_5}{3\sqrt{2}}, \quad x_3' = -x_3,$$

$$x_4' = \frac{2x_1 + x_2 + 3x_4 - 2x_5}{3\sqrt{2}}, \quad x_5' = \frac{x_1 - 2x_2 + 2x_4 + 3x_5}{3\sqrt{2}}$$

RESOLUCION

La matriz ortogonal A de f en la base canónica de $\mathbb{R}^5$ y la matriz $B = (1/2)(A + A^t)$ son:

$$A = \begin{bmatrix} 3/3\sqrt{2} & -2/3\sqrt{2} & 0 & -2/3\sqrt{2} & -1/3\sqrt{2} \\ 2/3\sqrt{2} & 3/2\sqrt{2} & 0 & -1/3\sqrt{2} & 2/3\sqrt{2} \\ 0 & 0 & -1 & 0 & 0 \\ 2/3\sqrt{2} & 1/3\sqrt{2} & 0 & 3/3\sqrt{2} & -2/3\sqrt{2} \\ 1/3\sqrt{2} & -2/3\sqrt{2} & 0 & 2/3\sqrt{2} & 3/3\sqrt{2} \end{bmatrix}$$

$$B=\begin{bmatrix} 1/\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 1/\sqrt{2} \end{bmatrix}$$

La matriz simétrica B ha resultado ser diagonal, pero el -1 de su diagonal principal lo queremos en el lugar 1, 1, lo que se consigue cambiando el orden de los vectores $\bar{e}_1$, $\bar{e}_2$, $\bar{e}_3$, $\bar{e}_4$ y $\bar{e}_5$ de la base canónica, poniendo $\bar{e}_3$ en el primer lugar. La forma canónica C de f es:

$$C=\begin{bmatrix} \boxed{-1} & 0 & 0 & 0 & 0 \\ 0 & \boxed{1/\sqrt{2} & -1/\sqrt{2}} & 0 & 0 \\ 0 & \boxed{1/\sqrt{2} & 1/\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1/\sqrt{2} & -1/\sqrt{2}} \\ 0 & 0 & 0 & \boxed{1/\sqrt{2} & 1/\sqrt{2}} \end{bmatrix}$$

Una base $(\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_n, \bar{u}_5)$ de $\mathbb{R}^4$ en la que C es la matriz de f tiene por primer vector a $\bar{u}_1 = \bar{e}_3$, como ya se ha indicado. Los otros cuatro vectores son del subespacio $V = \mathcal{V}(\bar{e}_1, \bar{e}_2, \bar{e}_4, \bar{e}_5)$, suplementario ortogonal del $\mathcal{V}(\bar{e}_3)$. Para hallar dichos vectores, empecemos tomando un vector cualquiera de V, por ejemplo, el $\bar{e}_1$, y hallando su imagen

$$f(\bar{e}_1) = \frac{1}{3\sqrt{2}}(3\bar{e}_1 - 2\bar{e}_2 - 2\bar{e}_4 - \bar{e}_5)$$

como dos de los vectores, de los cuatro que se andan buscando, han de formar una base ortonormal de $\mathcal{V}(\bar{e}_1, f(\bar{e}_1))$, ellos pueden ser

$$\bar{u}_2 = \bar{e}_1 \quad \text{y} \quad \bar{u}_3 = \frac{1}{3}(2\bar{e}_2 + 2\bar{e}_4 + \bar{e}_5)$$

Hemos de tomar ahora un nuevo vector de V que sea ortogonal a los $\bar{u}_2$ y $\bar{u}_3$, por ejemplo, $\bar{v} = \bar{e}_2 - 2\bar{e}_4 + 2\bar{e}_5$, y hallar su imagen por f, que resulta ser $f(\bar{v}) = -9\bar{e}_2 + 9\bar{e}_4$; como los dos últimos vectores, $\bar{u}_4$ y $\bar{u}_5$, han de formar una base ortonormal de $\mathcal{V}(\bar{v}, f(\bar{v}))$, ellos pueden ser

$$\bar{u}_4 = \frac{1}{\sqrt{2}}(\bar{e}_2 - \bar{e}_4) \quad \text{y} \quad \bar{u}_5 = \frac{1}{3\sqrt{2}}(\bar{e}_2 + \bar{e}_4 - 4\bar{e}_5)$$

CAPITULO

10 Forma canónica de Jordan

Cuando hablábamos, en el capítulo anterior, sobre la diagonalización de un endomorfismo $f: V \to V$, de un espacio vectorial V (real o complejo) de dimensión finita n, tuvimos ocasión de comprobar que dicha diagonalización no siempre era posible. Dicho de otra forma, hay matrices cuadradas que no son diagonalizables por semejanza. Cuando de entre todas las matrices A asociadas al endomorfismo f, en las distintas bases de V, no hay ninguna que sea diagonal, se busca alguna de ellas que sea lo mas «sencilla» posible, que tenga cuantos más elementos nulos mejor. En el caso complejo, la matriz de Jordan de f da respuesta a este problema:

Si f tiene n autovalores, entonces de entre sus matrices asociadas hay una, la de Jordan; que tiene la siguiente forma: Fuera de su diagonal y de la primera paralela a ella (por encima), todos sus demás elementos son nulos; su diagonal está formada por los n autovalores de f y la primera paralela a ella tiene sus elementos valiendo 1 ó 0:

$$J = \begin{bmatrix} \lambda_1 & \varepsilon_1 & & & \\ & \lambda_2 & \varepsilon_2 & & O \\ & & \ddots & \ddots & \\ & O & & \lambda_{n-1} & \varepsilon_{n-1} \\ & & & & \lambda_n \end{bmatrix} \qquad \begin{array}{l} \lambda_i \text{ autovalor de } f \\ \varepsilon_i = 1 \ o \ \varepsilon_i = 0 \end{array}$$

Se dedica este capítulo al estudio de esta cuestión, probando la existencia de tal matriz y estudiando su forma, en función de la estructura interna del endomorfismo f o de la de una matriz A asociada a f.

En términos de matrices, lo anterior significa que si una matriz real o compleja A, de tamaño $n \times n$, tiene n autovalores, entonces existe una matriz regular P tal que $J = P^{-1}AP$, donde J es la referida matriz de Jordan. Nos ocuparemos de obtener, no sólo J, sino también la matriz de paso P.

En el caso $n = 3$, las matrices de Jordan que no son diagonales son, solamente, de uno de los tres tipos siguientes:

$$\begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}, \quad \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix}, \quad \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}$$

Obsérvese que, en estos tres casos, para que aparezca un $\varepsilon = 1$ es necesario (no suficiente) que los dos autovalores contiguos a dicho ε sean iguales entre sí. Esto acontece, también, en el caso general (n cualquiera).

El interés de los resultados que vamos a obtener aquí, en este capítulo 10, radica, fundamentalmente, en su utilización en el estudio y búsqueda de soluciones de los sistemas de ecuaciones diferenciales lineales con coeficientes constantes.

α FORMA DE JORDAN DE UNA MATRIZ

10.1. DIAGONALIZACION EN BLOQUES TRIANGULARES

Con independencia del interés que estas cuestiones tienen por sí mismas, van a sernos aquí de utilidad, en el estudio de las formas de Jordan, del que constituyen el primer paso.

SUBESPACIOS INVARIANTES DE LOS ENDOMORFISMOS

[186]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial V, y sea U un subespacio de V. Se dice que U es *subespacio invariante* por f si todo vector $\bar{u} \in U$ se transforma, por f, en un vector $f(\bar{u})$ que también es de U; esto es, si $f(U) \subset U$. Si U es invariante, la restricción $f|_U$, de f a U, es un endomorfismo de U en U, que denotaremos poniendo $f: U \to U$.

EJEMPLOS

1. Si λ es un autovalor del endomorfismo $f: V \to V$, entonces el subespacio propio V_λ, de f, asociado al autovalor λ, es un subespacio invariante por f ya que, para todo $\bar{u} \in V_\lambda$ es $f(\bar{u}) = \lambda\bar{u}$, que también pertenece a V_λ.
2. Sea $f: \mathbb{R}^3 \to \mathbb{R}^3$ la rotación de ángulo α, cualquiera, alrededor de una recta dada. El endomorfismo f tiene como subespacio invariante al plano (subespacio vectorial bidimensional) ortogonal a la recta que es eje de la rotación.
3. Sea $f: \mathbb{R}^4 \to \mathbb{R}^4$ el endomorfismo que, respecto de cierta base $(\bar{e}_1, \bar{e}_2, \bar{e}_3, \bar{e}_n)$ tiene por matriz a la

$$A = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 7 & -3 & 4 & 5 \\ 2 & 1 & -2 & -5 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Los subespacios $\mathcal{V}(\bar{e}_2, \bar{e}_3)$ y $\mathcal{V}(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ son invariantes por f.

DIAGONALIZACION EN BLOQUES

[187]

Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$ de dimensión finita, y sean $U_1, U_2, ..., U_p$ subespacios de V invariantes por f tales que $V = U_1 \oplus U_2 \oplus \cdots \oplus U_p$. Si A_i es la matriz del endomorfismo $f: U_i \to U_i$ respecto de cierta base B_i de U_i, para $i = 1, 2, ..., p$, entonces la matriz de $f: V \to V$ en la base $B = B_1 \cup B_2 \cup \cdots \cup B_p$ es la siguiente matriz A, diagonal por bloques:

$$A = \begin{bmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots \\ O & O & \cdots & A_p \end{bmatrix}$$

En tal caso, se dice que f es un endomorfismo diagonalizable en bloques)(*).

(*) Para indicar que, con estas hipótesis, la matriz A es la señalada, suele decirse que A es la suma directa de las matrices A_i y se pone $A = A_1 \oplus A_2 \oplus \cdots \oplus A_p$. Llamando f_i al endomorfismo $f: U_i \to U_i$ (restricción de f al subespacio invariante U_i) y en el supuesto de que $V = U_1 \oplus U_2 \oplus \cdots \oplus U_p$, se pone $f = f_1 \oplus f_2 \oplus \cdots \oplus f_p$ y se dice que f es la suma directa de los f_i.

DEMOSTRACION

La comprobación del anterior aserto es inmediata: basta con tener en cuenta que las columnas de la matriz A y de las matrices A_i están formadas por las coordenadas en la base B y en las bases B_i, respectivamente, de las imágenes de los vectores de una y otras bases.

EJEMPLO

El endomorfismo $f: \mathbb{R}^3 \to \mathbb{R}^3$ que transforma (x, y, z) en $(2x, -2x - z, 3x + 3y + 3z)$ es diagonalizable por bloques ya que, llamando $\bar{a} = (1, -1, 0)$, $\bar{b} = (0, 1, -1)$ y $\bar{c} = (0, 0, 1)$, como $f(\bar{a}) = 2\bar{a}$, $f(\bar{b}) = \bar{b} + \bar{c}$ y $f(\bar{c}) = -b + 2\bar{c}$, la matriz de f en la base $(\bar{a}, \bar{b}, \bar{c})$ es la:

$$\left[\begin{array}{c|cc} 2 & 0 & 0 \\ \hline 0 & 1 & -1 \\ 0 & 1 & 2 \end{array}\right]$$

ENDOMORFISMOS Y MATRICES TRIANGULARIZABLES

Vamos a dar un nuevo paso en la dirección que lleva a la forma canónica de Jordan. Dado un endomorfismo f no diagonalizable, se va a estudiar si existe alguna matriz de f que sea triangular; de ser así se dirá que f es triangularizable.

[188]

1.ª Sea $f: V \to V$ un endomorfismo, del espacio vectorial $V \neq O$ de dimensión finita n; sea K el cuerpo de escalares. Se dice que f es *triangularizable* si existe una base de V en la que f tiene matriz T triangular. Se verifica que f es triangularizable si y sólo si f tiene n autovalores, en K(*). En este caso, la diagonal de T está formada por los autovalores de f (cada uno aparece tantas veces como indique su multiplicidad algebraica).

2.ª Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K. Se dice que A es triangularizable (por semejanza) si existe una matriz triangular T que es semejante con A. Se verifica que A es triangularizable si y sólo si A tiene n autovalores, en K (*). En este caso, la diagonal de T está formada por los autovalores de f (figurando cada uno tantas veces como indique su multiplicidad algebraica).

(*) Si $K = \mathbb{C}$ esta condición se verifica siempre.

DEMOSTRACION

Nótese en primer lugar que si f tiene matriz triangular superior en una base $(\bar{e}_1, \bar{e}_2, .., \bar{e}_{n-1}, \bar{e}_n)$ de V, entonces la matriz de f en la base $(\bar{e}_n, \bar{e}_{n-1}, ..., \bar{e}_2, \bar{e}_1)$ es triangular inferior. Por tanto, basta con considerar el caso de matrices triangulares superiores.

La triangularización de un endomorfismo es, obviamente, equivalente a la triangularización (por semejanza) de su matriz asociada en cualquier base de V. Las proposiciones 1.ª y 2.ª son, pues, equivalentes. Podemos suponer, y así lo haremos, que A es la matriz de f en una cierta base de V.

Comprobemos primero que si f es triangularizable, es decir, si f tiene matriz triangular $T = [t_{ij}]$ en cierta base de V, entonces f tiene n autovalores. Así ocurre, ya que el polinomio característico de f, que es el $\det(T - \lambda I)$, como T es triangular, resulta ser:

$$\det(T - \lambda I) = (t_{11} - \lambda)(t_{22} - \lambda)...(t_{nn} - \lambda)$$

que tiene por raíces a $\lambda = t_{11}$, $\lambda = t_{22}$, ..., $\lambda = t_{nn}$; así, pues, f tiene por autovalores a los n elementos de la diagonal de T.

Supongamos ahora que f (y A) tiene n autovalores (valores iguales o distintos en K). Para probar que, entonces, f es triangularizable, procedamos por inducción sobre n. Esta propiedad es evidentemente cierta para $n = 1$ (toda matriz 1×1 es triangular). Suponiendo, pues, que la propiedad es cierta siempre que sea $\dim V = n - 1$ o lo que es igual para las matrices que tienen tamaño $(n-1) \times (n-1)$, hemos de probar que también es cierta si $\dim V = n$ o si A tiene tamaño $n \times n$.

Como se supone que f tiene n autovalores, entonces tiene uno al menos, que llamaremos λ_1. Si $\bar{u}_1 \neq \bar{o}$ es un vector propio de f asociado al autovalor λ_1,

la matriz A' de f en una base $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$, cuyo primer vector sea $\bar{u}_1$, tiene su primera columna formada por los elementos $\lambda_1, 0, ..., 0$, lo que nos permite expresarla en la forma

$$A' = \left[\begin{array}{c|ccc} \lambda_1 & * & * \cdots & * \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right]$$

A_1 = matriz de tamaño $(n-1) \times (n-1)$.
$*$ denota a un cierto escalar de K (el valor que pueda tener $*$ carece de interés en nuestros razonamientos).

El polinomio característico de f, que tiene n raíces, es entonces

$$\det (A' - \lambda I) = (\lambda_1 - \lambda) \det (A_1 - \lambda I)$$

luego el polinomio característico de A_1 tiene $n-1$ autovalores. Por tanto, según la hipótesis de inducción, A_1 es triangularizable (por semejanza), es decir, existe una matriz regular P_1, de tamaño $(n-1) \times (n-1)$, tal que $T_1 = P_1^{-1} A_1 P_1$ es triangular. Sea P la siguiente matriz, que es regular y tiene, evidentemente, por inversa a la P^{-1} que a continuación se señala:

$$P = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & P_1 & \\ 0 & & & \end{array}\right] \qquad P^{-1} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & P_1^{-1} & \\ 0 & & & \end{array}\right]$$

Recurriendo a P y P^{-1}, al realizar el producto $P^{-1}A'P$, se obtiene fácilmente que:

$$P^{-1}A'P = \left[\begin{array}{c|ccc} \lambda_1 & * & * \cdots & * \\ \hline 0 & & & \\ \vdots & & P_1^{-1}A_1P_1 & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} \lambda_1 & * & * \cdots & * \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right] = T$$

que es una matriz triangular superior. Como esta matriz T es semejante con A', resulta que T es matriz de f en alguna base de V, con lo que concluye la demostración.

EJEMPLO

La matriz

$$A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 2 & 1 \\ -1 & -1 & 1 \end{bmatrix}$$

es triangularizable, pues su ecuación característica es $(\lambda - 2)^3 = 0$, que tiene todas sus raíces reales ($\lambda = 2$ triple). Como el subespacio propio de A está formado por las matrices columna proporcionales a la siguiente matriz C_1, completando esta matriz C_1 con otras dos, C_2 y C_3, hasta formar una matriz P regular, la matriz $A' = P^{-1}AP$ tendrá ceros debajo de su elemento de lugar 1,1:

$$C_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad C_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad C_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad P = [C_1 \mid C_2 \mid C_3]$$

$$A' = P^{-1}AP = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & -1 & 1 \end{bmatrix} = \left[\begin{array}{c|cc} 2 & 1 & 1 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{A_1}} \\ 0 & & \end{array}\right] \quad \text{con} \quad A_1 = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}$$

Se comprueba fácilmente que la matriz A_1 tiene el subespacio propio (correspondiente al autovalor $\lambda = 2$) formado por las matrices columna proporcionales a la siguiente matriz D_1; completando ésta con otra matriz, D_2, se obtiene una matriz regular Q_1 tal que $A_1 = Q_1^{-1}A_1Q_1$ es ya triangular y nos permite obtener una forma triangular T de la matriz dada A:

$$D_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad Q_2 = [D_1 \mid D_2], \quad Q = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{Q_2}} \\ 0 & & \end{array}\right]$$

$$T = Q^{-1}A'Q = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Nótese que $T = R^{-1}AR$, donde $R = PQ$ es la matriz de cambio.

EJERCICIO (TRIANGULARIZACION ORTOGONAL)

Sea $f: V \to V$ un endomorfismo, del espacio vectorial euclídeo $V \neq O$ de dimensión finita n. Se dice que f es triangularizable ortogonalmente si existe una base ortonormal de V en la que f tiene matriz T triangular.

Pruébese que f es triangularizable ortogonalmente si y sólo si f tiene sus n autovalores reales. En este caso, la diagonal de T está formada por los autovalores de f.

RESOLUCION

Para comprobar este ejercicio no hay más que seguir el mismo camino utilizado para probar la anterior proposición [188]. Lo allí dicho sigue siendo ahora de aplicación si se toma alguna precaución adicional: La base $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n)$ se

tomará ortonormal y la hipótesis de inducción deberá decir ahora que la triangularización, que se supone para el caso $n-1$, lo es ortogonal, o dicho de otra forma, se supondrá que la matriz P_1 es ortogonal; de ser así, es obvio que la matriz P es entonces ortogonal, lo que prueba la triangularización ortogonal.

OBSERVACION

Sea A una matriz, de tamaño $n \times n$, que tiene n autovalores, es decir, que tiene polinomio característico del tipo:

$$\det (A - \lambda I) = (-1)^n(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_p)^{m_p}$$

donde

$$m_1 + m_2 + \cdots + m_p = n$$

Según se ha probado anteriormente (véase [188]), la matriz A es entonces semejante a una matriz triangular superior T cuya diagonal está formada por los autovalores de A o, más exactamente, es la sucesión de los autovalores:

$$\lambda_1, \overset{m_1}{\ldots}, \lambda_1, \lambda_2, \overset{m_2}{\ldots}, \lambda_2, \ldots, \lambda_p, \overset{m_p}{\ldots}, \lambda_p$$

Una tal matriz triangular podría ser, por ejemplo, la siguiente, que se ha dividido en bloques de manera que cada uno de sus bloques diagonales tenga iguales todos los elementos de la diagonal:

$$\left[\begin{array}{c|cc|cc} 2 & 7 & 5 & 1 & 3 \\ \hline 0 & 1 & 4 & 2 & 3 \\ 0 & 0 & 1 & 6 & 1 \\ \hline 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 & 4 \end{array}\right]$$

Es evidente que, en general, la matriz triangular T puede expresarse, en bloques, de la siguiente forma:

$$T = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1p} \\ O & T_{22} & \cdots & T_{2p} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ O & O & \cdots & T_{pp} \end{bmatrix} \qquad [1]$$

T_{ij} (para $i \leqslant j$) bloque de tamaño $m_i \times m_j$
T_{ii} bloque triangular superior que tiene por diagonal a la sucesión $\lambda_i, \lambda_i, \ldots, \lambda_i$

Pues bien, se puede conseguir que esta matriz triangular, semejante a A, tenga nulos todos los bloques no diagonales, o sea, que $T_{ij} = O$ para $i \neq j$. Esto es:

DIAGONALIZACION EN BLOQUES TRIANGULARES

[189]

Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K, que tiene n autovalores en K y sean $\lambda_1, \lambda_2, ..., \lambda_p$ los autovalores distintos de A, que tienen multiplicidades algebraicas $m_1, m_2, ..., m_p$ (con $m_1 + m_2 + \cdots + m_p = n$). Se verifica entonces que A es semejante a una matriz del tipo:

$$\begin{bmatrix} T_1 & O & \cdots & O \\ O & T_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ O & O & \cdots & T_p \end{bmatrix} \qquad [2]$$

donde T_i es un bloque triangular superior, de tamaño $m_i \times m_i$, cuya diagonal es la sucesión $\lambda_i, \lambda_i, \overset{m_i}{\ldots}, \lambda_i$ (para $i = 1, 2, ..., p$).

DEMOSTRACION

Según se acaba de recordar, A es semejante a una matriz como la matriz T de la página anterior (la señalada con [1] y, por ello, bastará con comprobar que T es semejante a una matriz del tipo [2], descrita en el enunciado. Vamos, pues, a buscar una matriz regular P tal que $P^{-1}TP$ sea del tipo [2], y dicha búsqueda se hará recurriendo a las matrices elementales (de tamaño $n \times n$) como la

$$E(hk, a) = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & \cdots & a & \\ & & & \ddots & \vdots & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow h \\ \\ \leftarrow k \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & \uparrow & & \uparrow & \\ & & h & & k & \end{matrix}$$

que se obtiene, a partir de la matriz unidad, cambiando el 0 de lugar hk (con $h \neq k$) por el escalar $a \in K$. Nótese que la matriz $E(hk, a)$ es regular (para cualquiera que sea $a \in K$) y que su inversa es $E(hk, -a)$, con lo que, para cualquiera que sea la matriz M, de tamaño $n \times n$, se verifica que:

$$E(hk, a)\, M\, E(hk, -a) \quad \text{es semejante a } M$$

Obsérvese también que, para cualquier matriz M, se verifica:

- $E(hk, a) \cdot M$ es la matriz que resulta de realizar en M la siguiente manipulación: a la fila h-ésima se le suma la fila k-ésima multiplicada por a.
- $M \cdot E(hk, -a)$ es la matriz que resulta de realizar en M la siguiente manipulación: a la columna k-ésima se le suma la columna h-ésima multiplicada por $-a$.

Por tanto, fijados unos h y k con $h < k$, la matriz

$$T' = E(hk, a)\, T\, E(hk, -a)$$

que es semejante a T (triangular superior), difiere de T en los elementos de la fila h-ésima situados a la derecha del hk, en los elementos de la columna k-ésima situados por encima del hk y en el propio elemento hk que ahora vale

$$t'_{hk} = t_{hk} - a t_{hh} + a t_{kk} = t_{hk} - a\delta_{hk}, \quad \text{con} \quad \delta_{hk} = t_{hh} - t_{kk}$$

Si los lugares h y k se toman de manera que correspondan a dos bloques distintos de los T_{ii}, en que descompusimos T (véase [1] de la página anterior), resulta que t_{hh} y t_{kk} son dos autovalores distintos de A y su diferencia δ_{hk} es, pues, no nula. Tomando entonces $a = t_{hk} / \delta_{hk}$, se consigue que sea $t'_{hk} = 0$.

Vamos a proceder sucesivamente del modo que se acaba de indicar: de manera que, en cada paso, sera $h < k$, h y k corresponderán a distintos bloques T_{ii} y se tomará $a = t_{hk} : (t_{hh} - t_{kk})$. Y lo haremos de manera que hk recorra, uno tras otro, todos los bloques T_{ij} con $i \neq j$ y éstos los recorra siguiendo el siguiente orden:

$$\begin{array}{cccc} i, j = p-1, p & i, j = p-2, p-1 & i, j = p-2, p & \cdots \\ ij = 1, 2 & ij = 1, 3 & \cdots & ij = 1, p \end{array}$$

(de abajo a arriba y de izquierda a derecha); dentro de cada bloque, también se avanzará de abajo a arriba y de izquierda a derecha. Siguiendo este camino, los bloques T_{ij} (con $i \neq j$) se van transformando, sucesivamente, en bloques nulos, según es fácil comprobar. Al acabar el proceso, se habrá obtenido, pues, una matriz semejante con T, y por tanto, semejante con A, que es del tipo [2], que se define en el enunciado, como había que comprobar.

10.2. FORMA CANONICA DE JORDAN

MATRICES Y BLOQUES DE JORDAN

Según se acaba de comprobar (véase [189]), toda matriz cuadrada $n \times n$ que tenga n autovalores es semejante a una matriz diagonal por bloques, cuyos bloques diagonales son matrices triangulares T_i de tamaño $m_i \times m_i$; cada una de estas matrices T_i, tiene todos los elementos de su diagonal iguales a un auto-

valor λ_i de A, cuya multiplicidad algebraica es m_i. Pues bien, se pretende comprobar que es posible conseguir (transformando A por semejanza) que estos bloques T_i sean del tipo:

$$J(\lambda_i) = \begin{bmatrix} \lambda_i & \varepsilon_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & \varepsilon_2 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_i & \varepsilon_{m-1} \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix} \quad \begin{array}{l} \text{donde } \varepsilon_s = 0 \text{ o } \varepsilon_s = 1 \\ (\text{para } s = 1, 2, \ldots, m-1) \end{array}$$

Interesa señalar que una matriz triangular de este tipo es, a su vez, diagonal por bloques triangulares J_j del siguiente tipo, que se llaman *cajas o bloques de Jordan*:

$$J(\lambda) = \begin{bmatrix} J_1 & O & \cdots & O \\ O & J_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots \\ O & O & \cdots & J_q \end{bmatrix} \quad \text{donde } J_j \text{ es:}$$

$$J_j = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} \quad (\text{para } j = 1, 2, \ldots, q) \qquad [1]$$

Obsérvese que por cada ε_s que es nulo se produce una partición de $J(\lambda_i)$ en bloques; si hay q cajas de Jordan, es que hay $q-1$ de los $n-1$ valores ε_s que son nulos

$$J_j \rightarrow \left[\begin{array}{ccc} \ddots & \ddots & \\ \lambda_i & 1 & \\ 0 & \lambda_i & 0 \\ & 0 & \lambda_i \quad 1 \\ & & 0 \quad \lambda_i \quad \ddots \\ & & \quad\quad\quad \ddots \end{array}\right] \leftarrow J_{j+1}$$

Resumiendo, lo que pretendemos comprobar es que toda matriz A, de tamaño $n \times n$, que tenga n autovalores λ_i (iguales o distintos), es semejante a una *matriz de Jordan* J, que es diagonal por bloques $J(\lambda_i)$, cada uno de los

cuales es, a su vez, diagonal por bloques, siendo estos últimos de los llamados bloques de Jordan, que acabamos de describir en [1]. Cada bloque $J(\lambda_i)$ corresponde a un autovalor λ_i de A; todos los elementos de la diagonal de $J(\lambda_i)$ valen λ_i; si m_i es la multiplicidad algebraica de λ_i, entonces $J(\lambda_i)$ tiene tamaño $m_i \times m_i$; si d_i es la multiplicidad geométrica de λ_i, entonces $J(\lambda_i)$ consta de d_i cajas de Jordan.

EJEMPLOS

1. La siguiente matriz J es una matriz de Jordan, que consta de los tres bloques de Jordan J_1, J_2 y J_3; los dos primeros corresponden al autovalor $\lambda = 2$ (triple) y el tercero al autovalor $\lambda = 5$ (triple):

$$J = \left[\begin{array}{cc|c|ccc} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right] \quad J_1 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad J_2 = [2] \quad J_3 = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

2. Respecto de los elementos situados sobre la diagonal de una matriz de Jordan, que valen $\varepsilon = 1$ o $\varepsilon = 0$, conviene observar que si se cambian de lugar (un $\varepsilon = 1$ y un $\varepsilon = 0$), la nueva matriz de Jordan que así se obtiene no será semejante a la anterior, y ello aunque el cambio de producto entre elementos ε situados sobre autovalores iguales. Así, las siguientes matrices de Jordan J y J' no son semejantes:

$$J = \left[\begin{array}{cccc|c} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ \hline 0 & 0 & 0 & 0 & 3 \end{array}\right] \qquad J = \left[\begin{array}{ccc|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ \hline 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right]$$

3. Unas cajas de Jordan especialmente interesantes son las correspondientes a autovalores $\lambda = 0$. Una tal caja N_q, de tamaño $q \times q$, es tal que su potencia q-ésima es nula, $(N_q)^q = O$. Así, por ejemplo:

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad N^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$N^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad N^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = O$$

Para llegar al resultado que se ha comentado anteriormente, que se conoce como teorema de Jordan, vamos a proceder por etapas. Empezaremos suponiendo que todos los autovalores son iguales; más tarde se abordará el caso general. Antes de demostrar el teorema, comprobemos la siguiente propiedad:

PROPOSICION

[190] Sea M una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K, que expresada en bloques es de la forma:

$$M = \begin{bmatrix} 0 & A_1 & A_2 & \cdots & A_q \\ O & N_1 & O & \cdots & O \\ O & O & N_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & & N_q \end{bmatrix}, \quad \text{donde} \quad N_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

O es un bloque nulo.
$0 \in K$ elemento nulo.
N_i de tamaño $h_i \times h_i$
A_i de tamaño $1 \times h_i$ $\quad (h_1 + h_2 + \cdots + h_q = n)$

Se verifica que la matriz M es semejante a la siguiente matriz M', de igual tamaño e igualmente descompuesta en bloques que M:

$$M' = \begin{bmatrix} 0 & \Delta & O & \cdots & O \\ O & N_1 & O & \cdots & O \\ O & O & N_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & & N_q \end{bmatrix}$$

donde Δ (de tamaño $1 \times h_1$) es $\Delta = [\delta\ 0\ 0\ \cdots\ 0]$, siendo $\delta = 1$ o $\delta = 0$.

DEMOSTRACION

Recurriendo a ciertas matrices regulares P_1, P_2, P_3, P_4, que iremos dando a continuación, se obtendrán sucesivas matrices semejantes con M;

$$M_1 = P_1^{-1}MP_1 \quad , \quad M_2 = P_2^{-1}M_1P_2 \quad , \quad M_3 = P_3^{-1}M_2P_3 \quad , \quad M_4 = P_4^{-1}M_3P_4$$

hasta conseguir que la última, M_4, sea la matriz M' de enunciado. Veamos quiénes son éstas P_i y M_i:

1. Llamando B_i a la matriz, de igual tamaño que A_i, que queda definida por:

$$\text{si } A_i = [a, b, c, ..., f, g], \quad \text{entonces} \quad B_i = [b, c, d, ..., g, 0]$$

y tomando para P_1 la siguiente matriz (cuya inversa es, obviamente, la P_1^{-1} que también se da), se llega a la $M_1 = P_1^{-1}MP_1$ que se indica:

$$P_1 = \begin{bmatrix} 1 & -B_1 & -B_2 & \cdots & -B_q \\ O & I_1 & O & \cdots & O \\ O & O & I_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & \cdots & I_q \end{bmatrix}$$

$$P_1^{-1} = \begin{bmatrix} 1 & B_1 & B_2 & \cdots & B_q \\ O & I_1 & O & \cdots & O \\ O & O & I_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & \cdots & I_q \end{bmatrix}$$

$$M_1 = \begin{bmatrix} 0 & \Delta_1 & \Delta_2 & \cdots & \Delta_q \\ O & N_1 & O & \cdots & O \\ O & O & N_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & \cdots & N_q \end{bmatrix}$$

donde Δ_i = matriz fila $(1 \times h_i)$ cuyos elementos son, salvo quizá el primero de ellos, todos los demás nulos, esto es:

$$\Delta_i = [\delta_i, 0, 0, ..., 0]$$

2. Si las anteriores matrices Δ_i son todas nulas, la proposición está ya demostrada, pues la matriz M_1 es, entonces, la matriz M' del enunciado. Supondremos, pues, que hay alguna de las Δ_i que tiene su primer

elemento no nulo, $\delta_i \neq 0$. Pues bien, si este elemento $\delta_i \neq 0$ está situado en la columna r-ésima, recurriendo a la matriz

$$P_2 = \begin{bmatrix} 1 & & & & & & \\ & 0 & \cdots & \cdots & \cdots & 1 & \\ & \vdots & 1 & & & \vdots & \\ & \vdots & & \ddots & & \vdots & \\ & \vdots & & & 1 & \vdots & \\ & 1 & \cdots & \cdots & \cdots & 0 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{bmatrix}$$

(P_2 es la matriz que resulta de cambiar en la matriz unidad las filas segunda y r-ésima. Nótese que $P_2^{-1} = P_2$) nos encontramos con que la matriz $M_2 = P_2^{-1} M_1 P_2$ es la que resulta de cambiar entre sí, en M_1, las filas segunda y r-ésima y, también, las columnas segunda y r-ésima. En consecuencia, M_2 sólo difiere de M_1 en que se han cambiado de lugar los elementos δ_1 y δ_i. Así, pues, la matriz M_2 es de igual tipo que M_1 pero tiene la particularidad de que su elemento de lugar (1, 2) es no nulo.

3. Obtengamos ahora una nueva matriz $M_3 = P_3^{-1} M_2 P_3$, semejante a M, de igual tipo que M_2 pero en la que $\Delta_2 = O$, $\Delta_3 = O$, ..., $\Delta_q = O$. Esta se consigue tomando la siguiente matriz P_3, cuya inversa P_3^{-1} es, obviamente, la que se indica

$$P_3 = \begin{bmatrix} 1 & O & O & O & \cdots & O \\ O & I_1 & -D_2 & -D_3 & \cdots & -D_q \\ O & O & I_2 & O & \cdots & O \\ O & O & O & O & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & O & \cdots & I_q \end{bmatrix}$$

$$P_3^{-1} = \begin{bmatrix} 1 & O & O & O & \cdots & O \\ O & I_1 & D_2 & D_3 & \cdots & D_q \\ O & O & I_2 & O & \cdots & O \\ O & O & O & I_3 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ O & O & O & O & & I_q \end{bmatrix}$$

donde D_i es (para $i = 2, \ldots, q$) la matriz, de tamaño $h_1 \times h_i$, que tiene en la diagonal todos sus elementos valiendo $d_i = \delta_i / \delta_1$ y fuera de la diagonal todos sus elementos son nulos.

Obsérvese que la matriz $M_3 = P_3^{-1}M_2P_3$ es, efectivamente del tipo descrito, ya que

$$M_3 = \begin{bmatrix} 0 & \Delta_1 & \Delta_2' & \Delta_3' & \cdots & \Delta_q' \\ O & N_1 & D_2' & D_3' & \cdots & D_q' \\ O & O & N_2 & O & \cdots & O \\ O & O & O & N_3 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & O & & N_q \end{bmatrix}$$

donde (para $i = 2, ..., q$):

$$\Delta_i' = \Delta_i - \Delta_1 - \Delta_1 D_i = \Delta_i - \Delta_i = O$$
$$D_i' = D_i N_i - N_1 D_i = O$$
$$\Delta_1 = [\delta_1, 0, 0, ..., 0],\ \delta_1 \neq 0$$

4. Finalmente, echando mano de la siguiente matriz regular P_4 (cuya inversa P_4^{-1} es la que se señala), se obtiene una matriz $M_4 = P_4^{-1}M_3P_4$, semejante a M, que es evidentemente la matriz M' del enunciado con $\delta = 1$:

$$P_4 = \begin{bmatrix} \delta_1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}, \quad P_4^{-1} = \begin{bmatrix} 1/\delta_1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & 1 \end{bmatrix}$$

$$M_4 = \begin{bmatrix} 0 & \Delta & O & \cdots & O \\ O & N_1 & O & \cdots & O \\ O & O & N_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ O & O & O & \cdots & N_q \end{bmatrix}$$

donde $\Delta = [1, 0, 0, ..., 0]$.

Obsérvese que aquí, al final, se ha llegado a que $\Delta = [1, 0, ..., 0]$; ello no contradice a lo dicho en el enunciado, en el que se dice que puede ser:

$$\Delta = [1, 0, ..., 0] \quad \text{ó} \quad \Delta = [0, 0, ..., 0] = O$$

Esto es así, ya que, en el apartado segundo de la demostración, se presentaron las dos posibilidades $\Delta \neq O$ y $\Delta = O$; el caso $\Delta = O$ ya cumplía las exigencias del enunciado, por lo que no se siguió con él; el caso $\Delta \neq O$ es el que ha conducido a $\Delta = [1, 0, ..., 0]$.

FORMA DE JORDAN (CASO DE AUTOVALORES IGUALES)

[191]

Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K. Si A tiene un solo autovalor $\lambda \in K$ con multiplicidad algebraica n, entonces A es semejante a una matriz diagonal por bloques de Jordan, del siguiente tipo $J(\lambda)$; es decir, existe una matriz regular P tal que $P^{-1}AP = J(\lambda)$, siendo:

$$J(\lambda) = \begin{bmatrix} J_1 & O & \cdots & O \\ O & J_2 & \cdots & O \\ \cdots & \cdots & \cdots & \cdots \\ O & O & \cdots & J_q \end{bmatrix} \quad \text{donde} \quad J_j = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

El número q, de bloques de Jordan J_j ($j = 1, 2, ..., q$) que contiene $J(\lambda)$, es igual a la multiplicidad geométrica del autovalor (único) λ de A (*).

(*) Respecto de los tamaños de los q bloques de Jordan J_j y de la matriz P, véanse los apartados [193] y [194].

DEMOSTRACION

Recurramos al método de inducción, que lo aplicaremos sobre n. Para $n = 1$, la matriz A es, ella misma, un bloque de Jordan (de tamaño 1×1), por lo que el teorema se verifica obviamente. Suponiendo, pues, que el teorema es cierto para matrices A de tamaño $(n-1) \times (n-1)$, hemos de comprobar que también lo es para matrices de tamaño $n \times n$.

La matriz A, como tiene n autovalores en K (todos iguales), es triangularizable (véase [188], 2), o sea, es semejante a una matriz triangular T cuyos elementos de la diagonal son todos iguales a λ. Esta matriz T puede entonces expresarse en la forma:

$$T = \left[\begin{array}{c|c} \lambda & A_0 \\ \hline O & T' \end{array}\right]$$

donde T' es triangular de tamaño $(n-1) \times (n-1)$ y tiene todos los elementos de su diagonal iguales a λ. La matriz A_0 de tamaño $1 \times (n-1)$ no tiene ninguna particularidad.

De acuerdo con la hipótesis de inducción y como T' es una matriz de tamaño $(n-1) \times (n-1)$ y tiene sus $n-1$ autovalores iguales a λ, la matriz T' es semejante a una matriz diagonal por bloques de Jordan $J'(\lambda)$, de modo que

existe una matriz regular Q' de tamaño $(n-1) \times (n-1)$ tal que $J'(\lambda) = Q'^{-1}T'Q'$. Por tanto, recurriendo a la siguiente matriz regular Q, de tamaño $n \times n$, cuya inversa es la Q^{-1} que se señala,

$$Q = \left[\begin{array}{c|c} 1 & O \\ \hline O & Q' \end{array}\right] \quad Q^{-1} = \left[\begin{array}{c|c} 1 & O \\ \hline O & Q'^{-1} \end{array}\right]$$

resulta evidente que, para cierta matriz A'_0 de tamaño $1 \times (n-1)$, es:

$$Q^{-1}TQ = \left[\begin{array}{c|c} \lambda & A'_0 \\ \hline O & J'(\lambda) \end{array}\right] = A'$$

Esta última matriz, que hemos llamado A', es pues semejante a A. Hay pues que probar que existe una matriz regular P, de tamaño $n \times n$, tal que $P^{-1}A'P$ es una matriz $J(\lambda)$, descrita en el enunciado. Ahora bien, la relación $J(\lambda) = P^{-1}A'P$ puede ponerse de las siguientes formas, equivalentes a ella, en las que I es la matriz unidad de tamaño $n \times n$:

$$J(\lambda) - \lambda I = P^{-1}A'P - \lambda I \quad \text{o sea} \quad J(\lambda) - \lambda I = P^{-1}(A' - \lambda I)P$$

Por consiguiente, todo se reduce a comprobar que la matriz $A' - \lambda I$ es semajante a una matriz del tipo $J(\lambda) - \lambda I$. Ahora bien, esto no es otra cosa que lo ya demostrado en la anterior proposición [190]. La matriz M, de allí, no es otra que la $A' - \lambda I$, como evidencia el hecho de que $J_i - \lambda I = N_i$ (aquí J_i es el bloque de Jordan que se detalla en el enunciado); téngase en cuenta que se ha denotado

$$A'_0 = [A_1\ A_2 \cdots A_q]$$

La matriz que en [190] se llamó M' es la $J(\lambda) - \lambda I$ de ahora; nótese a este respecto que los cuatro bloques de la parte superior izquierda de M' forman el bloque:

$$\left[\begin{array}{c|c} 0 & \Delta \\ \hline O & N_1 \end{array}\right] = \left[\begin{array}{c|ccccccc} 0 & \delta & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{array}\right]$$

y éste corresponde a una sola caja de Jordan si es $\delta = 1$ y corresponde a dos (separadas por la línea de puntos) si es $\delta = 0$.

Finalmente, comprobemos que el número de bloques de Jordan de $J(\lambda)$ es igual a la multiplicidad geométrica del autovalor λ. Como existe una matriz

regular P tal que $J(\lambda) = P^{-1}AP$, procediendo como antes se obtiene fácilmente que

$$J(\lambda) - \lambda I = P^{-1}(A - \lambda I)P$$

Por tanto $J(\lambda) - \lambda I$ y $A - \lambda I$ tienen el mismo rango, que llamaremos r. La multiplicidad geométrica del autovalor λ es $n - r$. La matriz $J(\lambda) - \lambda I$ tiene todos sus elementos nulos menos algunos situados en la paralela a su diagonal que valen 1; el número de elementos de esta paralela que valen 1 es $n - q$ (q número de bloques de $J(\lambda)$) y, por ello, el rango de $J(\lambda) - \lambda I$ es $n - q$. Por tanto, $r = n - q$, o sea q (número de bloques de Jordan) es igual a $n - r$ (multiplicidad geométrica de λ).

EJEMPLO

Aún cuando los resultados que se han obtenido hasta ahora no son completos, pues no se ha dicho nada sobre el tamaño de los bloques de Jordan ni sobre cómo determinar la matriz P, estamos en condiciones de abordar algún ejemplo sencillo:

Obtener la forma canónica de Jordan J de la siguiente matriz A. Determinar también una matriz regular P tal que sea $J = P^{-1}AP$:

$$A = \begin{bmatrix} 0 & -3 & -1 \\ 2 & 5 & 1 \\ -2 & -3 & 1 \end{bmatrix}$$

RESOLUCION

El polinomio característico de A es

$$\det(A - \lambda I) = -\lambda^3 + 6\lambda^2 - 12\lambda + 8 = -(\lambda - 2)^3$$

luego A tiene sus tres autovalores reales e iguales a $\lambda = 2$. Como

$$\text{rang}\,(A - 2I) = \text{rang}\begin{bmatrix} -2 & -3 & -1 \\ 2 & 3 & 1 \\ -2 & -3 & -1 \end{bmatrix} = 1$$

resulta que la multiplicidad geométrica del autovalor $\lambda = 2$ de A es $d = 3 - \text{rang}\,(A - 2I) = 2$. Así pues, la forma canónica J tiene dos bloques de Jordan, lo que obliga a que sea

$$J = \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 2 \end{array}\right]$$

Para hallar una matriz P tal que $J = P^{-1}AP$, poniendo esta igualdad en la forma

$PJ = AP$ y expresando P en función de sus tres columnas, o sea, poniendo $P = [C_1 \mid C_2 \mid C_3]$, se tiene que:

$$[C_1 \mid C_2 \mid C_3]J = A[C_1 \mid C_2 \mid C_3]$$

o sea,

$$AC_1 = 2C_1, \quad AC_2 = 2C_2 + C_1, \quad AC_3 = 2C_3$$

que podemos poner en la forma

$$(A - 2I)C_1 = O, \quad (A - 2I)C_2 = C_1, \quad (A - 2\lambda)C_3 = O$$

Por tanto, C_1 y C_3 son dos autovectores, linealmente independientes, de A; esto es, C_1 y C_3 son soluciones de $(A - 2I)C = O$, o sea:

$$\begin{bmatrix} -2 & -3 & -1 \\ 2 & 3 & 1 \\ -2 & -3 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

luego

$$C_1, C_3 = \begin{bmatrix} \alpha \\ \beta \\ -2\alpha - 3\beta \end{bmatrix} \quad (\alpha, \beta \in \mathbb{R}) \qquad [1]$$

La ecuación que afecta a C_2 será:

$$\begin{bmatrix} -2 & -3 & -1 \\ 2 & 3 & 1 \\ -2 & -3 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ -2\alpha - 3\beta \end{bmatrix}$$

que tiene solución si $\alpha + \beta = 0$; tomando $\alpha = 1$ y $\beta = -1$, una solución para C_2 es la:

$$C_2 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \text{ correspondiente a } \quad C_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Tomando, finalmente, para C_3 una autovector de los obtenidos en [1] que sea linealmente independiente de C_1, por ejemplo, tomando $\alpha = 1$ y $\beta = 0$, se obtiene:

$$P = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 0 \\ 1 & -1 & -2 \end{bmatrix}$$

FORMA DE JORDAN (CASO DE AUTOVALORES DISTINTOS)

Apoyándonos en la diagonalización por bloques triangulares, que obtuvimos en [189], y en la forma de Jordan para el caso de autovalores iguales, vamos a obtener la forma de Jordan en el caso general:

[192]

Sea A una matriz cuadrada de tamaño $n \times n$ cuyos elementos pertenecen a un cuerpo K, que tiene n autovalores en K. Llamando $\lambda_1, ..., \lambda_p$ a los autovalores distintos de A y si $m_1, ..., m_p$ son sus multiplicidades algebraicas ($m_1 + \cdots + m_p = n$), entonces A es semejante a una matriz J de Jordan, es decir, $J = P^{-1}AP$ para cierta matriz regular P, donde J es una matriz del siguiente tipo:

$$J = \begin{bmatrix} J(\lambda_1) & & & O \\ & J(\lambda_2) & & \\ & & \ddots & \\ O & & & J(\lambda_p) \end{bmatrix} \quad \text{con} \quad J(\lambda_i) = \begin{bmatrix} J_{i1} & & & O \\ & J_{i2} & & \\ & & \ddots & \\ O & & & J_{iq_i} \end{bmatrix} \quad (i = 1, ..., p)$$

donde $J(\lambda_i)$ tiene tamaño $m_i \times m_i$ y los J_{ij} (para $j = 1, 2, ..., q_i$) son bloques de Jordan que tienen, todos ellos, los elementos de la diagonal iguales al autovalor λ_i, es decir:

$$J_{ij} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}$$

El número q_i de bloque de Jordan J_{ij} ($j = 1, 2, ..., q_i$) que contiene $J(\lambda_i)$, es igual a la multiplicidad geométrica de autovalor λ_i de A (*).

(*) Respecto de los tamaños de los q_i bloques de Jordan J_{ij} y de la matriz regular P, véanse los apartados [193] y [194].

DEMOSTRACION

De acuerdo con el resultado que obtuvimos en [189], al estudiar la diagonalización por bloques triangulares, y como A tiene n autovalores, se verifica que A es semejante a una matriz del tipo

$$T = \begin{bmatrix} T_1 & & O \\ & T_2 & \\ O & & \ddots \\ & & T_p \end{bmatrix}$$

donde T_i (para $i = 1, 2, ..., p$) es una matriz triangular superior de tamaño $m_i \times m_i$ cuya diagonal tiene todos sus elementos iguales al autovalor λ_i.

A cada una de las matrices T_i ($i = 1, 2, ..., p$) les es de aplicación lo dicho en el apartado anterior (véase [191]), sobre la forma de Jordan para matrices con todos sus autovalores iguales; por tanto, T_i es semejante a una matriz del tipo $J(\lambda_i)$ descrita en el enunciado, es decir, existe (para cada $i = 1, 2, ..., p$) una matriz regular P_i, de tamaño $m_i \times m_i$, tal que $J(\lambda_i) = P_i^{-1} T_i P_i$. Vamos a echar mano de la siguiente matriz P, que es regular y tiene como inversa a la P^{-1} que se indica:

$$P = \begin{bmatrix} P_1 & & & \\ & P_2 & O & \\ & O & \ddots & \\ & & & P_q \end{bmatrix} \quad P^{-1} = \begin{bmatrix} P_1^{-1} & & & \\ & P_2^{-1} & & O \\ & & \ddots & \\ & O & & P_q^{-1} \end{bmatrix}$$

De acuerdo con lo anterior, la matriz $P^{-1}TP$ será

$$P^{-1}TP = \begin{bmatrix} P_1^{-1}T_1P_1 & & & O \\ & P_2^{-1}T_2P_2 & & \\ & & \ddots & \\ & O & & P_q^{-1}T_2P_q \end{bmatrix} = \begin{bmatrix} J(\lambda_1) & & & O \\ & J(\lambda_2) & & \\ & & \ddots & \\ & O & & J(\lambda_p) \end{bmatrix} = J$$

luego T es semejante a J y, como A y T eran semejantes, resulta a la postre que A y J son semejantes, como había que comprobar.

EJEMPLO

A pesar de que, como ocurría con el ejemplo anterior, aún está sin estudiar cómo se determinan, en un caso general, los tamaños de las cajas de Jordan de una matriz A y el modo de localizar una matriz P que permite poner la forma de Jordan J como $J = P^{-1}AP$, ello no impide que podamos abordar ya estas cuestiones en algún caso sencillo, como el siguiente:

Obtener la forma canónica de Jordan, J, de la siguiente matriz A. Determinar también una matriz regular P tal que $J = P^{-1}AP$:

$$A = \begin{bmatrix} -4 & 0 & -4 & -1 \\ -4 & 3 & 0 & 4 \\ 7 & 0 & 7 & 1 \\ -3 & 0 & -4 & -2 \end{bmatrix}$$

RESOLUCION

El polinomio característico de A es

$$\det(A - \lambda J) = (\lambda - 3)^2(\lambda + 1)^2$$

luego A tiene sus cuatro autovalores reales: $\lambda = 3$ doble y $\lambda = -1$ doble. Como

$$\text{rang}\,(A - 3I) = 2 \quad \text{y} \quad \text{rang}\,(A + 1I) = 3,$$

las multiplicidades geométricas de los autovalores $\lambda = 3$ y $\lambda = -1$ son 2 y 1, respectivamente. Todo lo anterior conduce, obviamente, a que la forma de Jordan de A es la matriz:

$$J = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Para hallar una matriz P tal que $J = P^{-1}AP$, expresando esta relación en la forma $PJ = AP$ y recurriendo a las columnas de P, es decir, poniendo $P = [C_1 \mid C_2 \mid C_3 \mid C_4]$, se obtiene que:

$$AC_1 = 3C_1 \quad , \quad AC_2 = 3C_2 \quad , \quad AC_3 = -C_3 \quad , \quad AC_4 = -C_4 + C_3 \qquad [1]$$

Por tanto, C_1 y C_2 son dos autovectores, linealmente independientes, de A correspondientes al autovalor $\lambda = 3$; esto es C_1 y C_2 son soluciones, linealmente independientes, de $(A - 3I)C = O$, o sea de:

$$\begin{bmatrix} -7 & 0 & -4 & -1 \\ -4 & 0 & 0 & 4 \\ 7 & 0 & 4 & 1 \\ -3 & 0 & -4 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

luego,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ -2\alpha \\ \alpha \end{bmatrix} \quad (\text{para } \alpha, \beta \in \mathbb{R})$$

Tomando, por ejemplo, $\alpha = 0$ y $\beta = 1$ se obtiene C_1 y para $\alpha = 1$ y $\beta = 0$ obtenemos C_2.

La columna C_3 es, de acuerdo con la tercera de las ecuaciones [1], un autovector de A correspondiente a $\lambda = -1$; esto es, C_3 es solución no nula de $(A + 1I)C = O$, o sea de:

$$\begin{bmatrix} -3 & 0 & -4 & -1 \\ -4 & 4 & 0 & 4 \\ 7 & 0 & 8 & 1 \\ -3 & 0 & -4 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

luego

$$C_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

Finalmente, de la última de las ecuaciones [1], se obtiene:

$$\begin{bmatrix} -3 & 0 & -4 & -1 \\ -4 & 4 & 0 & 4 \\ 7 & 0 & 8 & 1 \\ -3 & 0 & -4 & -1 \end{bmatrix} C_4 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

luego

$$C_4 = \begin{bmatrix} 1-\lambda \\ 1 \\ \lambda - 1 \\ -\lambda \end{bmatrix} \quad (\text{para } \lambda \in \mathbb{R})$$

y tomando $\lambda = 1$, por ejemplo, se obtiene un C_4 independiente de las otras columnas.

Hemos, pues, obtenido:

$$P = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & -2 & -1 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

Conviene destacar lo laboriosa que ha resultado la búsqueda de las columnas de la matriz de cambio P; estas columnas, además de verificar a las relaciones [1], han de ser linealmente independientes, lo que obliga a tomar un sin fin de precauciones. Se entiende, pues, que el estudio que vamos a abordar ahora, encaminado a localizar la referida matriz P, en un caso general, sea más tortuoso de lo deseable y esté lleno de recovecos.

β OBTENCION DE LA FORMA DE JORDAN

En relación con la forma canónica de Jordan, hay algunas cuestiones que no hemos resuelto todavía, aun cuando en los casos sencillos se puedan atajar con facilidad. Nos referimos al modo de determinar los tamaños de los bloques de Jordan y a cómo hallar la matriz regular P que, por semejanza, transforma una matriz dada A en su forma de Jordan J, esto es. aquella P que permite poner $J = P^{-1}AP$.

10.3. TAMAÑO DE LOS BLOQUES DE JORDAN

[193]

Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K, que tiene n autovalores en K y sean $\lambda_1, ..., \lambda_p$ los autovalores distintos de A, cuyas multiplicidades algebracias son $m_1, ..., m_p$ (con $m_1 + \cdots + m_p = n$). Según sabemos (véase [192]), la matriz canónica de Jordan, J, de A es de la forma:

$$J = \begin{bmatrix} J(\lambda_1) & & O \\ & J(\lambda_2) & \\ & & \ddots \\ O & & J(\lambda_p) \end{bmatrix}$$

donde

$$J(\lambda) = \begin{bmatrix} J_1 & & O \\ & J_2 & \\ & & \ddots \\ O & & J_q \end{bmatrix} \quad (\text{para } \lambda = \lambda_1, ..., \lambda_p)$$

(La matriz $J(\lambda)$, de tamaño $m \times m$, tiene todos los elementos de su diagonal iguales a λ; las matrices $J_1, J_2, ..., J_q$ son bloques de Jordan.)

En este supuesto, se verifica que, para cada $\lambda = \lambda_1, ..., \lambda_p$, existe un $s \in \mathbb{N}$ tal que:

$$\text{rang } (A - \lambda I) \geqslant \text{rang } (A - \lambda I)^2 \geqslant \cdots \geqslant \text{rang } (A - \lambda I)^s = n - m \quad [1]$$

y llamando k_j (para $j = 1, 2, ..., s$) a los números naturales

$$k_1 = n - \text{rang}\,(A - \lambda I) \quad \text{y} \quad k_j = \text{rang}\,(A - \lambda I)^{j-1} - \text{rang}\,(A - \lambda I)^j, \text{ si } j > 1$$

entonces el número h_j de bloques de Jordan de tamaño $j \times j$ que contiene $J(\lambda)$ es:

$$h_j = k_j - k_{j+1} \text{ (para } j = 1, 2, ..., s-1) \quad \text{y} \quad h_s = k_s \quad [2]$$

DEMOSTRACION

Como A y J son semejantes, existe una matriz regular P tal que $J = P^{-1}AP$, lo que permite poner:

$$J - \lambda I = P^{-1}AP - \lambda I = P^{-1}AP - P^{-1}(\lambda I)P = P^{-1}(A - \lambda I)P$$

de donde resulta que, por ser P regular, $A - \lambda I$ y $J - \lambda I$ tienen el mismo rango.

Por otra parte, al ser distintos los elementos de la diagonal de cada una de las matrices $J(\lambda_1), J(\lambda_2), ..., J(\lambda_p)$, resulta evidente que

$$n - \text{rang } (J - \lambda I) = m - \text{rang } (J(\lambda) - \lambda I)$$

y que

$$n - \text{rang}\,(J - \lambda I)^j = m - \text{rang}\,(J(\lambda) - \lambda I)^j \quad j = 2, \ldots, s$$

Nótese que, si J_j es un bloque de Jordan de tamaño $j \times j$ y que tiene los elementos de su diagonal valiendo λ, las matrices

$$J_j - \lambda I,\ (J_j - \lambda I)^2,\ \ldots,\ (J_j - \lambda I)^{j-1},\ (J_j - \lambda I)^j = O$$

tienen todos sus elementos nulos excepto, respectivamente, los de la 1.ª, 2.ª, ... paralela a su diagonal, que valen 1; por tanto, es:

$$\text{rang}\,(J_j - \lambda I) = j - 1,\ \text{rang}\,(J_j - \lambda I)^2 = j - 2,\ \ldots,\ \text{rang}\,(J_j - \lambda I)^{j-1} = 1$$

Ahora bien, llamando $h_1, h_2, \ldots, h_s$ a los números de bloques de Jordan que tiene $J(\lambda)$ de tamaños 1, 2, ..., s, respectivamente, de acuerdo con lo que se acaba de decir, se verifica que:

$$\begin{aligned}
m - \text{rang}\,(J(\lambda) - \lambda I) &= h_1 + h_2 + h_3 + \cdots + h_s = \\
&= m - (0h_1 + 1h_2 + 2h_3 + \cdots + (s-1)h_s) \\
m - \text{rang}\,(J(\lambda) - \lambda I)^2 &= m - \quad (0h_2 + 1h_3 + \cdots + (s-2)h_s) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
m - \text{rang}\,(J(\lambda) - \lambda I)^{s-1} &= m - \quad (1h_s) \\
m - \text{rang}\,(J(\lambda) - \lambda I)^s &= m
\end{aligned}$$

Resulta entonces que para los rangos de las matrices $(A - \lambda I)^j$ es:

$$\begin{aligned}
n - \text{rang}\,(A - \lambda I) &= m - (h_2 + 2h_3 + \cdots + (s-1)h_s) = h_1 + h_2 + \cdots + h_s \\
n - \text{rang}\,(A - \lambda - \lambda I)^2 &= m - (h_3 + 2h_4 + \cdots + (s-2)h_s) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
n - \text{rang}\,(A - \lambda I)^{s-1} &= m - (h_s) \\
n - \text{rang}\,(A - \lambda I)^s &= m
\end{aligned}$$

Así, pues, se verifican las relaciones [1] del enunciado y el número s es el mayor de los tamaños de los bloques de Jordan de que consta $J(\lambda)$. Los números k_j que se definen en el enunciado resultan valer. a la vista de lo anterior:

$$\begin{aligned}
k_1 &= h_1 + h_2 + h_3 + \cdots + h_s \\
k_2 &= h_2 + h_3 + \cdots + h_s \\
k_3 &= h_3 + \cdots + h_s \\
&\cdots\cdots\cdots\cdots \\
&\cdots\cdots\cdots\cdots \\
k_{s-1} &= h_{s-1} + h_s \\
k_s &= h_s
\end{aligned}
\qquad \text{luego} \qquad
\begin{cases}
h_1 = k_1 - k_2 \\
h_2 = k_2 - k_3 \\
h_3 = k_3 - k_4 \\
\cdots\cdots\cdots\cdots \\
\cdots\cdots\cdots\cdots \\
h_{s-1} = k_{s-1} - k_s \\
h_s = k_s
\end{cases}$$

con lo que concluye la demostración.

EJERCICIO

Hallar la matriz de Jordan, J, de la siguiente matriz A:

$$A = \begin{bmatrix} 4 & 0 & 1 & 0 \\ 1 & 1 & 1 & -1 \\ 0 & -2 & 3 & -1 \\ -1 & 2 & -1 & 4 \end{bmatrix}$$

RESOLUCION

El polinomio característico de A es:

$$\det (A - \lambda I) = \begin{vmatrix} 4-\lambda & 0 & 1 & 0 \\ 1 & 1-\lambda & 1 & -1 \\ 0 & -2 & 3-\lambda & -1 \\ -1 & 2 & -1 & 4-\lambda \end{vmatrix} = (\lambda - 3)^4$$

Hay, pues, un solo autovalor $\lambda = 3$ con multiplicidad algebraica $m = 4$. Todas las cajas de Jordan de J tienen entonces su diagonal ocupada por $\lambda = 5$. Para hallar los tamaños de estas cajas, procediendo como se señala en la anterior proposición [193], calculemos las sucesivas potencias de $A - 3I$:

$$A - 3I = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & -2 & 0 & -1 \\ -1 & 2 & -1 & 1 \end{bmatrix}$$

$$(A - 3I)^2 = \begin{bmatrix} 1 & -2 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (A - 3I)^3 = O$$

Por tanto, los valores de los k_j y h_j son, en este caso:

$$k_1 = 4 - \text{rang } (A - 3I) = 4 - 2 = 2, \qquad h_1 = k_1 - k_2 = 1$$
$$k_2 = \text{rang } (A - 3I) - \text{rang } (A - 3I)^2 = 2 - 1 = 1, \qquad h_2 = k_2 - k_3 = 0$$
$$k_3 = \text{rang } (A - 3I)^2 - \text{rang } (A - 3I)^3 = 1 - 0 = 1, \qquad h_3 = k_3 = 1$$

de donde resulta que J tiene una caja 1×1 y otra caja 3×3. En consecuencia, J es:

$$J = \left[\begin{array}{ccc|c} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ \hline 0 & 0 & 0 & 3 \end{array}\right]$$

10.4. MATRIZ DE PASO A LA FORMA DE JORDAN

SUBESPACIOS PROPIOS GENERALIZADOS

Los subespacios propios de una matriz $n \times n$ o de su endomorfismo asociado, de K^n en K^n, son subespacios invariantes independientes, esto es, con suma directa. Pero, salvo si la matriz es diagonalizable (por semejanza), esta suma no llega a ser todo K^n. Pues bien, asociados a cualquier matriz $n \times n$ que tenga n autovalores en K (iguales o distintos), hay ciertos subespacios invariantes, cada uno de ellos asociado a un autovalor distinto de la matriz, que incluyen a los correspondientes subespacios propios y que se llaman subespacios propios generalizados de la matriz:

PROPOSICION

[194]

Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K, que tiene n autovalores en K y sean $\lambda_1, \lambda_2, ..., \lambda_p$ los autovalores distintos de A, cuyas multiplicidades algebraicas son $m_1, m_2, ..., m_p$ (con $m_1 + m_2 + \cdots + m_p = n$). Se llaman *subespacios propios generalizados* de A a los:

$$V_i = \{X \in \mathcal{M}_{n\times 1} / (A - \lambda_i)^{m_i} X = O\}, \quad \dim V_i = m_i \text{ (para } i = 1, 2, ..., p)$$

Se verifica que estos subespacios son invariantes por el endomorfismo $X \rightarrowtail AX$ y tales que

$$\mathcal{M}_{n\times 1} = V_1 \oplus V_2 \oplus \cdots \oplus V_p$$

Si J es la matriz de Jordan de A y P es una matriz regular tal que $J = P^{-1}AP$, entonces: sus m_1 primeras columnas forman base de V_1, sus m_2 siguientes columnas forman base de V_2, ..., sus m_p últimas columnas forman base de V_p.

DEMOSTRACION

Recurriendo a la matriz de Jordan J y a la matriz P que permite poner $J = P^{-1}AP$, se obtiene que (para $i = 1, 2, ..., p$):

$$P^{-1}(A - \lambda_i I)P = P^{-1}AP - \lambda_i I = J - \lambda_i I$$

Elevando a la potencia m_i el primer y último miembros de la anterior igualdad, resulta que:

$$\begin{aligned}(J - \lambda_i I)^{m_i} &= [P^{-1}(A - \lambda_i I)P][P^{-1}(A - \lambda_i I)P] \overset{m_i}{\cdots} [P^{-1}(A - \lambda_i I)P] = \\ &= P^{-1}(A - \lambda_i I)^{m_i}P\end{aligned} \qquad [1]$$

Consideremos el endomorfismo $f_i: \mathcal{M}_{n\times 1} \to \mathcal{M}_{n\times 1}$ que en la base usual tiene asociada la matriz $(A - \lambda_i I)^{m_i}$. El núcleo de f_i es el subespacio V_i.. Al referir f_i a la base de $\mathcal{M}_{n\times 1}$ formada por las columnas de la matriz P y de acuerdo con la igualdad [1], resulta que la matriz de f_i es ahora $(J - \lambda_i I)^{m_i}$. Analicemos cómo es esta matriz. De acuerdo con la expresión de la matriz J (véase por ejemplo [193]), será

$$J - \lambda_i I = \begin{bmatrix} \boxed{J(\lambda_1) - \lambda_i I} & & & O \\ & \boxed{J(\lambda_2) - \lambda_i I} & & \\ & & \ddots & \\ O & & & \boxed{J(\lambda_p) - \lambda_i I} \end{bmatrix}$$

donde la matriz $J(\lambda_j) - \lambda_i I$, cuyo tamaño es $m_j \times m_j$, tiene todos los elementos de la diagonal valiendo $\lambda_j - \lambda_i$, en la paralela a la diagonal (por encima) tiene sus elementos valiendo 1 ó 0 y todos sus demás elementos son nulos. Resulta entonces obvio que

$$(J - \lambda_i I)^{m_i} = \begin{bmatrix} J_1^{(i)} & & & O \\ & J_2^{(i)} & & \\ & & \ddots & \\ O & & & J_p^{(i)} \end{bmatrix}$$

donde,

$$J_j^{(i)} = [J(\lambda_j) - \lambda_i I]^{m_i} = \begin{cases} \text{si } i \neq j, & \text{una matriz regular de tamaño } m_j \times m_j, \\ & \text{triangular superior, cuyos elementos de} \\ & \text{la diagonal valen todos } (\lambda_j - \lambda_i)^{m_i} \\ \text{si } i = j, & \text{la matriz nula de tamaño } m_i \times m_i \end{cases}$$

De esto se infiere que V_i, por ser el núcleo de f_i, es el subespacio engendrado por las columnas, C_h, de P relativas al bloque nulo $J_i^{(i)}$ de la matriz $(J - \lambda_i I)^{m_i}$ de f_i, esto es, las m_i columnas situadas entre los lugares $h = m_1 + \cdots + m_{i-1} + 1$ y $h = m_1 + \cdots + m_{i-1} + m_i$. Repitiendo lo anterior para $i = 1, 2, ..., p$, nos en-

contramos con que: las m_1 primeras columnas de P forman una base de V_1, a la que llamaremos P_1, las m_2 siguientes columnas de P forman una base de V_2, a la que llamaremos P_2, ..., sus m_p últimas columnas forman una base de V_p, a la que llamaremos P_p:

$$V_1 = \mathcal{V}(P_1),\ V_2 = \mathcal{V}(P_2),\ \ldots,\ V_p = \mathcal{V}(P_p) \quad \text{con} \quad P[P_1, P_2, \ldots, P_p] \qquad [2]$$

Además, como $\mathcal{M}_{n\times 1} = \mathcal{V}(P)$, de [2] se infiere que

$$\mathcal{M}_{n\times 1} = V_1 \oplus V_2 \oplus \cdots \oplus V_p$$

Sea $f: \mathcal{M}_{n\times 1} \to \mathcal{M}_{n\times 1}$ al endomorfismo que en la base usual tiene asociada la matriz A. Como $J = P^{-1}AP$, en la base de $\mathcal{M}_{n\times 1}$ formada por las columnas de P, la nueva matriz de f es J. Como J está formada por bloques diagonales de tamaños $m_1, m_2, \ldots, m_p$, resulta que las columnas P_1 de P (las m_1 primeras), las columnas P_2 de P (las m_2 siguientes)..., las columnas P_p de P (las m_p últimas) se transforman por f en $m_1, m_2, \ldots, m_p$ columnas que, respectivamente, sólo depende de las $P_1, P_2, \ldots, P_p$. De esto se deduce que, por ser $V_i = \mathcal{V}(P_i)$, para $i = 1, 2, \ldots, P$, los subespacios V_i son invariantes por f, con lo que concluye la demostración.

PROPOSICION

[195]

Sea A una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K, que tiene n autovalores en K y sean $\lambda_1, \lambda_2, \ldots, \lambda_p$ los autovalores distintos de A, cuyas multiplicidades algebraicas son $m_1, m_2, \ldots, m_p$ (con $m_1 + m_2 + \cdots + m_p = n$). Sea J la matriz de Jordan de A:

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & O & \\ & & \ddots & \\ & O & & J_p \end{bmatrix}$$

donde J_i es una matriz de tamaño $m_i \times m_i$ que tiene todos los elementos de la diagonal valiendo λ_i, en la paralela a la diagonal (por encima) tiene sus elementos valiendo 1 ó 0 y todos sus demás elementos son nulos.

Sea P una matriz tal que $PJ = AP$ (no necesariamente regular) y sea $P = [P_1, P_2, \ldots, P_p]$ su descomposición en bloques de $m_1, m_2, \ldots, m_p$ columnas. Entonces se verifica que (para $i = 1, 2, \ldots, p$) las columnas de P_i pertenecen al subespacio propio generalizado correspondiente al autovalor λ_i.

DEMOSTRACION

Por ser $PJ = AP$ y debido a que J es diagonal por bloques, se verifica que $P_iJ_i = AP_i$. Llamando I e I_i a las matrices unidad de tamaños $n \times n$ y $m_i \times m_i$, es evidente que

$$P_i(J_i - \lambda_i I_i) = (A - \lambda_i I)P_i$$

Si premultiplicamos a ambos miembros de esta igualdad por la matriz $(A - \lambda_i I)^{m_i - 1}$, se obtiene:

$$(A - \lambda_i I)^{m_i} P_i = (A - \lambda_i I)^{m_i - 2}(A - \lambda_i I) P_i (J_i - \lambda_i I_i) = (A - \lambda_i I)^{m_i - 2} P_i (J_i - \lambda_i I_i)^2 =$$
$$= (A - \lambda_i I)^{m_i - 3}(A - \lambda_i I) P_i (J_i - \lambda_i I_i)^2 = (A - \lambda_i I)^{m_i - 3} P_i (J_i - \lambda_i I_i)^3 =$$
$$= \cdots = P_i (J_i - \lambda_i I_i)^{m_i} = P_i O = O$$

esto es, todas las columnas X de P_i verifican a la ecuación $(A - \lambda_i I)^{m_i} X = O$, que (véase [194]) es la ecuación del subespacio generalizado asociado a λ_i.

MATRIZ DE PASO A LA FORMA DE JORDAN

Si una matriz cuadrada A, de tamaño $n \times n$, tiene n autovalores, entonces A es semejante a una matriz de Jordan J, determinada por A; existe, pues, una matriz regular P tal que $J = P^{-1}AP$. Nos vamos a ocupar aquí de cómo hallar una tal matriz P, sabiendo ya que existe.

OBSERVACIONES

[196]

Sean $\lambda_1, \lambda_2, \ldots, \lambda_p$ los autovalores distintos de la matriz cuadrada A, de tamaño $n \times n$, y sean $m_1, m_2, \ldots, m_p$ sus respectivas multiplicidades algebraicas; se supone que $m_1 + m_2 + \cdots + m_p = n$, con lo que A admite matriz canónica de Jordan J:

$$J = \begin{bmatrix} J(\lambda_1) & & & O \\ & J(\lambda_2) & & \\ & & \ddots & \\ O & & & J(\lambda_p) \end{bmatrix}$$

donde

$$J(\lambda_i) = \begin{bmatrix} J_1 & & & O \\ & J_2 & & \\ & & \ddots & \\ O & & & J_q \end{bmatrix} \quad (i = 1, 2, \ldots, p)$$

siendo $J_1, J_2, \ldots, J_q$ cajas de Jordan que tienen sus diagonales ocupadas por el mismo autovalor λ_i. Esta matriz J es semejante a A, lo que significa que existe una matriz regular P tal que $J = P^{-1}AP$; esta igualdad la pondremos en la forma equivalente $PJ = AP$, con P regular, que nos a a resultar más manejable.

1.ª Vamos a ocuparnos de las columnas de P relativas a uno de sus autovalores λ_i, que son las que ocupan los lugares comprendidos entre el $m_1 + \cdots + m_{i-1} + 1$ y el $m_1 + \cdots + m_{i-1} + m_i$, ambos inclusive. Supongamos

que las cajas de Jordan $J_1, J_2, ..., J_q$, que componen $J(\lambda_i)$, tienen tamaños $t_1 \times t_1, t_2 \times t_2, ..., t_q \times t_q$ (nótese que $t_1 + t_2 + \cdots + t_q = m_i$) y denotemos a las respectivas columnas de P mediante

$$C_1^{(1)}, C_2^{(1)}, ..., C_{t_1}^{(1)}, C_1^{(2)}, C_2^{(2)}, ..., C_{t_2}^{(2)}, ..., C_1^{(q)}, C_2^{(q)}, ..., C_{t_q}^{(q)} \qquad [1]$$

Al expresar que PJ y AP, por ser iguales, tienen iguales sus columnas y, en particular, tienen iguales sus columnas de los m_i lugares relativos a λ_i, se obtiene que:

$$\left.\begin{array}{l} AC_1^{(1)} = \lambda_i C_1^{(1)},\ AC_2^{(1)} = \lambda_i C_2^{(1)} + C_1^{(1)},\ ...,\ AC_{t_1}^{(1)} = \lambda_i C_{t_1}^{(1)} + C_{t_1-1}^{(1)} \\ AC_1^{(2)} = \lambda_i C_1^{(2)},\ AC_2^{(2)} = \lambda_i C_2^{(2)} + C_1^{(2)},\ ...,\ AC_{t_2}^{(2)} = \lambda_i C_{t_2}^{(2)} + C_{t_2-1}^{(2)} \\ \cdots\cdots\cdots\

BUSQUEDA DE LA MATRIZ DE PASO A LA FORMA DE JORDAN

[197]

Sea A una matriz cuadrada, $n \times n$, cuyos elementos pertenecen a un cuerpo K, que tiene n autovalores en K y sean $\lambda_1, ..., \lambda_p$ los autovalores distintos de A, cuyas multiplicidades algebraicas son $m_1, ..., m_p$ (con $m_1 + \cdots + m_p = n$); sea J la matriz de Jordan de A y sea P una matriz regular tal que $J = P^{-1}AP$ (se sabe que J y P existen). Para localizar P basta con tener presente que:

- Cada autovalor λ_i tiene asociadas m_i columnas de P (las que van de la $m_1 + \cdots + m_{i-1} + 1$ a la $m_1 + \cdots + m_{i-1} + m_i$); reuniéndolas se obtiene P.
- Las m (para $m = m_1, ..., m_p$) columnas de P asociadas a λ (para $\lambda = \lambda_1, ..., \lambda_p$) se pueden obtener con el siguiente esquema, por etapas(*):

Etapa	Denominación genérica de una columna	Número de columnas (independientes)	Expresiones que definen a las columnas	
1	C_s	$k_s = h_s$ (indep.) $\rightarrow$	$C_s \in M_s$	[S]
2	C_{s-1}	k_{s-1} (indep.) $= \begin{cases} k_s \rightarrow \\ h_{s-1} \rightarrow \end{cases}$	$C'_{s-1} = (A - \lambda I)C_s \in M_{s-1}$ $C''_{s-1} \in M_{s-1}$	[S−1]
s − j + 1	C_j	k_j (indep.) $= \begin{cases} k_{j+1} \rightarrow \\ h_j \rightarrow \end{cases}$	$C'_j = (A - \lambda I)C_{j+1} \in M_j$ $C''_j \in M_j$	[j]
s − 1	C_2	k_2 (indep.) $= \begin{cases} k_3 \rightarrow \\ h_2 \rightarrow \end{cases}$	$C'_2 = (A - \lambda I)C_3 \in M_2$ $C''_2 \in M_2$	[2]
s	C_1	k_1 (indep.) $= \begin{cases} k_2 \rightarrow \\ h_1 \rightarrow \end{cases}$	$C'_1 = (A - \lambda I)C_2 \in N_1$ $C''_1 \in N_1$	[1]

$$N_s \supset N_{s-1} \supset \cdots \supset N_2 \supset N_1 \qquad M_j = \mathcal{V}(\{C'_j\}) \oplus \mathcal{V}(\{C''_j\})$$

$$N_j = N_{j-1} \oplus M_j \qquad \dim M_j = k_j$$

Se ha llamado: $N_j = \{C \in \mathcal{M}_{n\times 1} / (A - \lambda I)^j C = O\}$

M_j = subespacio de N_j que es suplementario de N_{j-1}

Para cada C''_j, de [j] ($j = s, ..., 2, 1$), las columnas

$$C''_j,\ C'_j = (A - \lambda I)C''_j,\ ...,\ C'_2 = (A - \lambda I)C'_3,\ C'_1 = (A - \lambda I)C'_2$$

ocuparán lugares consecutivos en P, situados en orden contrario al anterior; a ellas corresponde en P una caja de Jordan.

(*) La notación es la misma que en [193]; esto es: s = mayor tamaño de entre los de las cajas de Jordan correspondientes al autovalor λ; $k_1 = n - \text{rang}\,(A - \lambda I)$, $k_j = \text{rang}\,(A - \lambda I)^{j-1} - \text{rang}\,(A - \lambda I)^j$ si $j > 1$; $h_j = k_j - k_{j+1}$ (si $j < s$), $h_s = k_s$. Recuérdese que h_j es el número de bloques de Jordan relativos a λ que tienen tamaño $j \times j$.

DEMOSTRACION

Comencemos con unas observaciones encaminadas a señalar cuáles son las cuestiones que hemos de demostrar y cuáles otras son consecuencias de ellas.

Empezaremos comprobando que el método propuesto en el enunciado es compatible, es decir, conduce siempre a una matriz P regular. Luego hemos de ver que para la matriz P así obtenida se verifica que $J = P^{-1}AP$.

Respecto de la existencia de solución, para cada etapa j del proceso, probaremos que todas las k_j columnas C_j pueden tomarse independientes y que pertenecen al espacio M_j, de dimensión k_j. Para ello, veremos que las primeras k_{j+1} columnas C_j, esto es, las $C_j' = (A - \lambda I)C_{j+1}$, pertenecen a N_j, son linealmente independientes y engendran un subespacio que tiene intersección nula con el N_{j-1}, por lo que pertenecen a M_j y; una vez probado esto, la existencia de las k_j columnas C_j linealmente independientes está garantizada, pues los C_j'' serán h_j columnas de M_j que añadidas a las k_{j+1} columnas C_j' formen una base de M_j.

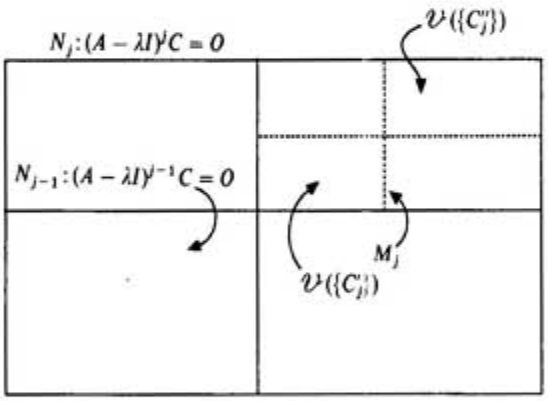


Figura 10.1

Al demostrar lo anterior, también queda de manifiesto que todas las columnas de las sucesivas etapas (correspondientes al autovalor λ de A) forman un sistema independiente (de $k_1 + k_2 + \cdots + k_s = m$ columnas), ya que no sólo son independientes las columnas de una misma etapa, como se acaba de hacer notar, sino que también lo son las de distintas etapas, ya que las columnas C_j, de una etapa, pertenecen a M_j y las C_i, de una etapa posterior, pertenecen a M_i y como $M_i \subset N_i \subset N_{j-1}$ y $M_j \cap N_{j-1} = O$, resulta que $M_i \cap M_j = O$.

Nótese que, si las m_i columnas relativas a λ_i (para $i = 1, 2, \ldots, p$) son independientes, entre sí, como ellas pertenecen al subespacio propio generalizado relativo a λ_i (véase [195]) y estos subespacios (para $i = 1, 2, \ldots, p$) son independientes (véase [194]), resulta que es independiente el conjunto de todas las n columnas de la matriz P la cual es, entonces, regular.

Así pues, se tiene:

1. Las k_{j+1} columnas $C_j' = (A - \lambda I)C_{j+1}$ pertenecen a N_j ya que, procedien-

do por inducción (para la primera etapa la propiedad es evidente), al suponer que $C_{j+1} \in N_{j+1}$, para la columna C'_j se obtiene:

$$(A - \lambda I)^j C'_j = (A - \lambda I)^j (A - \lambda I) C_{j+1} = (A - \lambda I)^{j+1} C_{j+1} = O$$

esto es C'_j pertenece a N_j.

2. Las k_{j+1} columnas $C'_j = (A - \lambda I) C_{j+1}$ son linealmente independientes. Recurriendo al método de inducción, como la propiedad es evidente para la primera etapa, partiremos de que las k_{j+2} columnas C'_{j+1} son independientes, con lo que lo son las k_{j+1} columnas C_{j+1}, y probaremos que, entonces, también lo son las k_j columnas C'_j. Supongamos, pues, que existen k_j escalares, que genéricamente denotaremos por α_j, tales que $\Sigma \alpha_j C'_j = O$; de ser así, se obtiene que

$$(A - \lambda I)^j (\Sigma \alpha_j C_{j+1}) = (A - \lambda I)^{j-1} [\Sigma \alpha_j (A - \lambda I) C_{j+1}] =$$
$$= (A - \lambda I)^{j-1} (\Sigma \alpha_j C'_j) = (A - \lambda I)^{j-1} O = O$$

Este resultado informa de que la columna $\Sigma \alpha_j C_{j+1}$ pertenece al subespacio N_j; ahora bien, esta columna $\Sigma \alpha_j C_{j+1}$ pertenece al subespacio M_{j+1}, pues a él pertenecen todas las columnas C_{j+1}, y como $N_j \cap M_{j+1} = O$, se ha de verificar que $\Sigma \alpha_j C_{j+1} = O$; como los C_{j+1} se han supuesto independientes, de la última igualdad se deduce que todos los α_j han de ser nulos, lo que prueba que los vectores C'_j son linealmente independientes.

3. El subespacio $\mathcal{V}(\{C'_j\})$, que engendran las k_{j+1} columnas C'_j, tienen intersección nula con el subespacio N_{j-1}. Para comprobar esto, bastaría con cerciorarse de que si, para ciertos escalares α_j, la columna $C = \Sigma \alpha_j C'_j$ pertenece a N_{j-1}, entonces es $C = O$. Así ocurre ya que:

$$C \in N_{j-1} \Rightarrow (A - \lambda I)^{j-1} (\Sigma \alpha_j C'_j) = O \Rightarrow$$
$$\Rightarrow (A - \lambda I)^{j-1} (\Sigma \alpha_j (A - \lambda I) C_{j+1}) = O \Rightarrow$$
$$\Rightarrow (A - \lambda I)^j (\Sigma \alpha_j C_{j+1}) = O \Rightarrow \Sigma \alpha_j C_{j+1} \in N_j$$

Como, igual que en el caso anterior, $\Sigma \alpha_j C_{j+1} \in M_{j+1}$ y como $N_j \cap M_{j+1} = O$, ha de ser $\Sigma \alpha_j C_{j+1} = O$. Ahora bien, por ser las columnas C_{j+1} independientes, de la última relación se deduce que todos los escalares α_j son nulos y, por ello, $C = \Sigma \alpha_i C'_j = O$, como había que comprobar.

4. Sea P una cualquiera de las matrices regulares que son solución del sistema planteado en el enunciado. las cuales existen según lo ya probado. Hemos de comprobar que, para tal matriz P, es $J = P^{-1}AP$. Esto quedará de manifiesto si verificamos que las columnas de PJ son respectivamente iguales a las de AP. Para ello fijémonos en las siguientes columnas de P que, según se dice en el enunciado, son consecutivas: a partir de una cualquiera de las columnas C''_j que satisfacen a la relación [j] (para cualquier $j = s, s-1, ..., 2, 1$), se consideran las C'_{j-1}, C'_{j-2}, ..., C'_2, C'_1 que derivan de ella, esto es las:

$$C'_1 = (A - \lambda I) C'_2,\ C'_2 = (A - \lambda I) C'_3,\ ....\ C'_{j-1} = (A - \lambda I) C''_j \quad \text{y} \quad C''_j \quad [*]$$

Veamos que $P^{-1}AP$ tiene, en las columnas de iguales lugares que las $C'_1, C'_2, ..., C'_{j-1}, C'_j{}'$, una caja de Jordan, de tamaño $j \times j$ y con diagonal ocupada por el autovalor λ. Esto es tanto como decir que son iguales las respectivas columnas, de dichos lugares, de AP y PJ, en el supuesto de que J tenga la referida caja de Jordan en el lugar indicado. Así ocurre, ya que las susodichas columnas de AP y PJ son:

$$\begin{array}{lllll} \text{en } AP: & AC'_1, & AC'_2, & AC'_3, & ..., \ AC'_j{}' \\ \text{en } PJ: & \lambda C'_1, & \lambda C'_2 + C'_1, & \lambda C'_3 + C'_2, & ..., \ \lambda C'_j{}' + C_{j-1} \end{array}$$

Y estas son respectivamente iguales, de acuerdo con la relación [j] del enunciado, a las [*] que acabamos de recordar.
Repitiendo este razonamiento con todas y cada una de las columnas $C'_j{}'$ $(j = s, s-2, ..., 2, 1)$, se obtiene que $P^{-1}AP$ es la pertinente matriz de Jordan.

EJERCICIO

Hallar una matriz regular P tal que $P^{-1}AP$ sea matriz de Jordan, donde

$$A = \begin{bmatrix} 4 & 0 & 1 & 0 \\ 1 & 1 & 1 & -1 \\ 0 & -2 & 3 & -1 \\ -1 & 2 & -1 & 4 \end{bmatrix}$$

RESOLUCION

Esta matriz ya fue considerada en el ejercicio que sigue a [193]. Allí vimos que A tenía un solo autovalor, $\lambda = 3$, que para él es (con la notación de [193] y [197]):

$$s = 3, \quad k_1 = 2, \quad k_2 = 1, \quad k_3 = 1, \quad h_1 = 1, \quad h_2 = 0 \quad \text{y} \quad h_3 = 1$$

y que la matriz de Jordan J de A es

$$J = \left[\begin{array}{ccc|c} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ \hline 0 & 0 & 0 & 3 \end{array}\right]$$

Para hallar P, siguiendo el camino marcado en [197], procediendo por etapas, se tiene:

- Los subespacios N_3, N_2 y N_1 tienen por ecuaciones a $(A - 3I)^3 C = O$,

$(A-3I)^2C=O$ y $(A-3I)C=O$; por tanto, los subespacios N_i son los formados por las columnas cuyos elementos c_1, c_2, c_3 y c_4 son:

$$N_3\colon (c_1, c_2, c_3, c_4) = \text{cualquiera}$$
$$N_2\colon c_1 - 2c_2 + c_3 - c_4 = 0$$
$$N_1\colon \begin{cases} c_1 + c_3 = 0 \\ 2c_2 + c_4 = 0 \end{cases} \Leftrightarrow (c_1, c_2, c_3, c_4) = (\alpha, \beta, -\alpha, -2\beta) \quad (\alpha, \beta \in \mathbb{R})$$

- Los subespacios M_3 y M_2 pueden ser los siguientes:

$$M_3 \oplus N_2 = N_3 \rightarrow M_3\colon (c_1, c_2, c_3, c_4) = (0, 0, 0, \lambda), \quad (\lambda \in \mathbb{R})$$
$$M_2 \oplus N_1 = N_2 \rightarrow M_2\colon (c_1, c_2, c_3, c_4) = (2h, h, 0, 0), \quad (h \in \mathbb{R})$$

- Las columnas de P pueden ser, entonces:

  1.ª etapa: hay una ($k_3 = 1$) columna $C_3 \in M_3$, que será (por razones tipográficas, escribimos la traspuesta C_3^t en lugar de dar C_3):

$$C_3^t = [0, 0, 0, 1]$$

  2.ª etapa: hay una ($k_3 = 1$) columna $C_2' = (A - 3I)C_3$ y ninguna ($h_2 = 0$) columna C_2'':

$$C_2''^t = [0, -1, -1, 1]$$

  3.ª etapa: hay una ($k_2 = 1$) columna $C_1' = (A - 3I)C_2$ y una ($h_1 = 1$) columna $C_1'' \in N_1$, que serán

$$C_1''^t = [-1, 0, 1, 0]$$
$$C_1'''^t = [0, 1, 0, -2]$$

En consecuencia, para la anterior matriz de Jordan J (ya obtenida en el ejercicio que sigue a [193]), la matriz P es:

$$P = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & -2 \end{bmatrix}$$

Ejercicios y problemas a la parte IV

ENUNCIADOS

IV.1. Sea V un espacio vectorial de dimensión finita y sean U_1 y U_2 dos subespacios suplementarios de V, de dimensiones p y q. Hallar los autovalores del endomorfismo, $f: V \to V$, proyección sobre U_1 paralelamente a U_2.

IV.2. Sea λ un autovalor de un endomorfismo $f: V \to V$. Compruébese que λ^2 es un autovalor del endomorfismo $f \circ f$. Si f es un automorfismo, pruébese que $1/\lambda$ es un autovalor de f^{-1}.

IV.3. Sean A y B dos matrices reales, ambas cuadradas y de igual tamaño. Compruébese que:

a) Si una, al menos, de las matrices A o B es regular, entonces AB y BA tienen el mismo polinomio característico.

b) Si A y B, ambas, singulares, entonces AB y BA tienen, también, el mismo polinomio característico (indicación: recurrir al resultado anterior aplicado a A y $B' = B + \varepsilon I$ y hállese luego que ε tiende a cero).

IV.4. Sea A una matriz cuadrada real de tamaño $n \times n$. Si n es par y $\det A < 0$, pruébese que A tiene, al menos, dos autovalores reales.

IV.5. Hallar los autovalores y los subespacios propios del endomorfismo $f: \mathbb{R}^3 \to \mathbb{R}^3$ que, en la base canónica, tiene asociada la matriz A:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 3 & -1 & 6 \\ 1 & -1 & 3 \end{bmatrix}$$

Analícese si f es diagonalizable.

IV.6. Hallar los autovalores y los subespacios propios del endomorfismo $f: \mathbb{R}^3 \to \mathbb{R}^3$ que, respecto de la base canónica, tiene asociada la siguiente matriz A:

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

Analícese si f es diagonalizable.

IV.7. Hallar los autovalores y los subespacios propios del endomorfismo $f: V \to V$ (donde V es un espacio vectorial real de dimensión 4) que, en cierta base dada $(\bar{e}_1, \bar{e}_2, \bar{e}_3, \bar{e}_4)$ de V, tiene asociada la siguiente matriz A:

$$A = \begin{bmatrix} 3 & 3 & 0 & 1 \\ -1 & -1 & 0 & -1 \\ 1 & 2 & 1 & 1 \\ 2 & 4 & 0 & 3 \end{bmatrix}$$

Analícese si f es diagonalizable.

IV.8. Hallar los autovalores y los subespacios propios de la matriz A, de tamaño $n \times n$ con $n \geqslant 2$:

$$A = \begin{bmatrix} 1+a & 1 & 1 & \cdots & 1 \\ 1 & 1+a & 1 & \cdots & 1 \\ 1 & 1 & 1+a & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1+a \end{bmatrix}$$

Analícese si A es diagonalizable por semejanza.

IV.9. Sea V un espacio vectorial real de dimensión 3, en el que se considera una base $B = (\bar{u}_1, \bar{u}_2, \bar{u}_3)$; en dicha base, las coordenadas se denotan por x_1, x_2, x_3. De un endomorfismo $f: V \to V$ se sabe que:

- El vector $6\bar{u}_1 + 2\bar{u}_2 + 5\bar{u}_3$ se transforma, por f, en sí mismo.
- $U = \{\bar{u} \in V / 2x_1 + 11x_2 - 7x_3 = 0\}$ es un subespacio propio de f.
- La traza de la matriz A de f, en B, es igual a 5.

a) Hallar los autovalores de f.

b) Hallar la matriz A de f en la base B.

IV.10. Hallar todas las matrices cuadradas, de un cierto tamaño $n \times n$, tales que $A = P^{-1}AP$ para toda matriz regular P, de tamaño $n \times n$.

IV.11. Dada una matriz cuadrada A, sea A' la matriz que resulta de permutar, en A, las filas i-ésima y j-ésima y, también, las columnas i-ésima y j-ésima. Analizar si A y A' son semejantes y, si lo son, hallar una matriz regular P tal que $A' = P^{-1}AP$.

IV.12. Sean A y A' dos matrices semejantes y sea P una matriz de paso, $A' = P^{-1}AP$. Caracterizar, en función de P, a todas las matrices Q de paso, esto es, tales que $A' = Q^{-1}AQ$.

IV.13. Comprobar que la siguiente matriz cuadrada A es diagonalizable en $\mathbb{C}$ y obtener su forma diagonal

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

IV.14. Se dice que una matriz cuadrada A es nilpotente si existe $k \in \mathbb{N}$ tal que $A^k = O$. Pruébese que una matriz cuadrada A, de tamaño $n \times n$, es nilpotente si y sólo si $\lambda = 0$ es el único autovalor de A, con multiplicidad n.

IV.15. Sea $f: V \to V$ un endomorfismo diagonalizable ($\dim V = n$) y sea $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_n)$ una base de V formada por vectores propios de f. Si es $\bar{u} = \bar{u}_1 + \bar{u}_2 + \cdots + \bar{u}_n$, compruébese que los vectores

$$\bar{e}_1 = \bar{u},\ \bar{e}_2 = f(\bar{u}),\ \bar{e}_3 = (f \circ f)(\bar{u}) = f^{(2)}(\bar{u}),\ \ldots,\ \bar{e}_n = f^{(n-1)}(\bar{u})$$

forman una base de V.

IV.16. Sea A la matriz real, dependiente del parámetro α:

$$A = \begin{bmatrix} 2\alpha + 4 & 1 - \alpha & -2\alpha - \alpha^2 \\ 0 & 4 - \alpha & 0 \\ 0 & 0 & 4 - \alpha^2 \end{bmatrix}$$

a) Obtener los valores de α para los que A es diagonalizable por semejanza.
b) Diagonalizar A para $\alpha = 1$ y para $\alpha = 2$.

IV.17. Sea $f: \mathbb{R}^3 \to \mathbb{R}^3$ el endomorfismo que, respecto de una base dada $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$, tiene asociada la matriz A:

$$A = \begin{bmatrix} 1 + \alpha & -\alpha & \alpha \\ 2 + \alpha & -\alpha & \alpha - 1 \\ 2 & -1 & 0 \end{bmatrix} \quad (\alpha \in \mathbb{R})$$

a) Obtener los autovalores de A, comprobando que no dependen de α.
b) Obtener los subespacios propios de f, en función de α, y estudiar si f es diagonalizable.
c) Cuando f sea diagonalizable, hallar su forma diagonal y la base correspondiente.

IV.18. Sea $f: \mathbb{R}^3 \to \mathbb{R}^3$ un endomorfismo del que se sabe lo siguiente:

- f es diagonalizable y sólo tiene dos autovalores distintos.
- $f(U) = V$, siendo

$$U = \{(x, y, z) \in \mathbb{R}^3 / x - 2y - z = 0\}$$

y

$$V = \mathbb{V}[(1, 0, 1), (-1, 1, 1)]$$

- $\lambda_1 = -1$ es un autovalor de f y uno de sus vectores propios pertenece a U.
- $(1, 0, -1)$ es un vector propio de f y está asociado a un autovalor simple.

a) Hallar la matriz A de f en la base canónica, en función de cuantos parámetros sea preciso.
b) Si en $\mathbb{R}^3$ se considera el producto escalar canónico, determinar f para que sea ortogonalmente diagonalizable.

IV.19. Sea A una matriz cuadrada de tamaño $n \times n$. Si A es antisimétrica, pruébese que sus autovalores son números complejos imaginarios puros o, bien, son nulos.

IV.20. En el espacio vectorial euclídeo canónico $\mathbb{R}^3$ y respecto de una base ortonormal $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$, se considera un endomorfismo del que se sabe que:

- $f(\bar{e}_1) = 3\bar{e}_1 + 2\bar{e}_2 + 2\bar{e}_3$, $f(\bar{e}_2) = 2\bar{e}_1 + 2\bar{e}_2$.
- La matriz A de f es simétrica.
- $\bar{a} = 2\bar{e}_1 - 2\bar{e}_2 - \bar{e}_3$ es autovector de f.

a) Hallar A y los autovalores y los autovectores de f.

b) Diagonalizar ortogonalmente f, determinando una base en la que se obtenga dicha diagonalización.

IV.21. En el espacio vectorial euclídeo canónico $\mathbb{R}^3$ y respecto de la base canónica, se considera el endomorfismo $f: \mathbb{R}^3 \to \mathbb{R}^3$ definido mediante

$$(x, y, z) \mapsto (x', y', z') \text{ siendo } \begin{cases} 3x' = -x + 2y + 2z \\ 3y' = 2x - y + 2z \\ 3z' = 2x + 2y - z \end{cases}$$

a) Razonar si f es ortogonalmente diagonalizable.
b) Razonar si f es una transformación ortogonal.
c) Hallar los autovalores y los subespacios propios de f.
d) Hallar la forma diagonal de f y una base ortonormal correspondiente.
e) Describir geométricamente la transformación f.

IV.22. Sea V un espacio vectorial euclídeo de dimensión 3 y sea $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ una base ortonormal de V. Considérese el endomorfismo $f: V \to V$ que en la base dada tiene asociada la matriz:

$$A = \begin{bmatrix} \alpha & \beta & 0 \\ \gamma & \alpha & \beta \\ 0 & \gamma & \alpha \end{bmatrix} \quad (\alpha, \beta, \gamma \in \mathbb{R})$$

a) Hallar la relación que debe haber entre los parámetros α, β y γ para que f admita un autovalor triple.
b) Suponiendo que $\beta\gamma > 0$, hallar los autovalores y subespacios propios de f (recúrrase al parámetro h, siendo $h^2 = 2\beta\gamma$).
c) Razónese si para algunos β y γ, la matriz A es ortogonalmente diagonalizable.

IV.23. De la siguiente matriz A se sabe que $\lambda_1 = 1$ es uno de sus autovalores y que $(1, 1, 1)$ es un vector propio de A asociado al autovalor λ_1:

$$A = \begin{bmatrix} 1 & 2 & \alpha \\ 2 & 1 & \beta \\ 2 & 2 & \gamma \end{bmatrix}$$

a) Hallar α, β y γ.
b) Hallar los autovalores y los subespacios propios de A.
c) Comprobar que A es diagonalizable y hallar su matriz diagonal D.
d) Analizar si A es ortogonalmente diagonalizable y, en caso afirmativo, hallar una matriz ortogonal P tal que $D = P^{-2}AP$.

IV.24. En el espacio vectorial euclídeo canónico $\mathbb{R}^3$, se considera un endomorfismo $f: \mathbb{R}^3 \to \mathbb{R}^3$ del que se sabe que:

- La matriz A de f en la base canónica es simétrica.
- El subespacio $V(2, -2, -1)$ es un subespacio propio de f.
- Los vectores $(1, 0, 0)$ y $(0, 1, 0)$ se transforman, respectivamente, en los vectores $(3, 2, 2)$ y $(2, 2, 0)$.

a) Razonar si f es ortogonalmente diagonalizable y si es una transformación ortogonal.
b) Hallar los autovalores y los subespacios propios de f.
c) Hallar la forma diagonal de f y obtener, si es posible, una base ortonormal de $\mathbb{R}^3$ en la que f tome dicha forma.

IV.25. Sea dada la siguiente matriz $A(\alpha)$:

$$A(\alpha) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1+\alpha & -1-\alpha \\ 1 & -1 & -1+\alpha & 1-\alpha \end{bmatrix}$$

($\alpha \in \mathbb{R}$ dado).

a) Hallar los autovalores de $A(\alpha)$.
b) Hallar los valores de α para los que $A(\alpha)$ es diagonalizable por semejanza.
c) Analizar si $A(\alpha)$ es ortogonalmente diagonalizable (con el producto escalar canónico) para algún valor α_0 de α y, si lo es, obtener la correspondiente diagonalización (esto es, la forma diagonal D y una matriz ortogonal de cambio, $D = P^{-1}A(\alpha_0)P$).
d) Describir geométricamente la transformación $f: \mathbb{R}^4 \to \mathbb{R}^4$ que, en la base canónica, tiene asociada la matriz $(1/2)A(\alpha_0)$.
e) Hallar el menor valor natural de h para el que la matriz $G = A(0) + hI$ es matriz métrica de un producto escalar.

IV.26. Sea $\omega: \mathbb{R}^3 \to \mathbb{R}$ la forma cuadrática dada por

$$\omega(x, y, z) = y^2 - 2xy - 2xz - 2yz$$

Diagonalizar ortogonalmente la forma cuadrática ω (*en* $\mathbb{R}^3$ se considera el producto escalar canónico).

IV.27. Sea $\omega: \mathbb{R}^3 \to \mathbb{R}$ la forma cuadrática que, en la base canónica de $\mathbb{R}^3$, tiene asociada la matriz simétrica:

$$A = \begin{bmatrix} 1 & -3 & -1 \\ -3 & 1 & 1 \\ -1 & 1 & 5 \end{bmatrix}$$

Hallar una base ortonormal de $\mathbb{R}^3$ (con el producto escalar canónico) en la que la matriz D de ω, que también se pide, sea diagonal.

IV.28. Sea A una matriz simétrica real y sean α y β el menor y el mayor de los autovalores de A. Llamando $A'(h) = A - hI$, hallar los valores de h para los que $A'(h)$ es definida positiva, definida negativa y no definida.

IV.29. Diagonalizar ortogonalmente la siguiente matriz simétrica A, hallando la correspondiente matriz diagonal D y una matriz de paso ortogonal P, esto es, tal que $P^{-1} = P^t$ y $D = P^{-1}AP$:

$$A = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix}$$

Analizar si A es definida positiva.

IV.30. Sean A y B dos matrices cuadradas simétricas, ambas de igual tamaño. Si los autovalores de A están en el intervalo $[a_1, a_2]$ y los autovalores de B están en el intervalo $[b_1, b_2]$ pruébese que los autovalores de $A + B$ están en el intervalo $[a_1 + b_1, a_2 + b_2]$.

IV.31. Sea V un espacio vectorial real de dimensión finita y sea $f: V \to V$ un endomorfismo. Obtener una condición necesaria y suficiente para que exista un producto escalar en V con el que f sea un endomorfismo simétrico.

IV.32. De una matriz cuadrada A de tamaño 6×6 se sabe que tiene los siguientes autovalores λ_1 y λ_2 y que éstos tienen las multiplicidades algebraicas (m_i) y geométricas (d_i) que se indican:

$$\lambda_1 = 3, \quad \text{con} \quad m_1 = 4 \quad \text{y} \quad d_1 = 2$$
$$\lambda_2 = 2, \quad \text{con} \quad m_2 = 2 \quad \text{y} \quad d_2 = 1$$

Obtener todas las posibles matrices de Jordan de A.

IV.33. De una matriz cuadrada A de tamaño 9×9 se sabe que tiene a $\lambda = 2$ como autovalor quíntuple y a $\lambda = -4$ como autovalor cuádruple. Se sabe además que

$$\text{rang}\,(A - 2I) = 7, \quad \text{rang}\,(A - 2I)^2 = 5,$$
$$\text{rang}\,(A - 2I)^3 = 4, \quad \text{rang}\,(A + 4I) = 8,$$
$$\text{rang}\,(A + 4I)^2 = 7, \quad \text{rang}\,(A + 4I)^3 = 6$$

Hallar las cajas de la matriz canónica de Jordan de A.

IV.34. Determinar, en función de los parámetros α, $\beta \in \mathbb{R}$, la forma canónica de Jordan de la matriz A

$$A = \begin{bmatrix} -1 & \alpha & 0 \\ 0 & -1 & \beta \\ 0 & 0 & 2 \end{bmatrix}$$

En el caso de que A no sea diagonalizable, hallar una matriz P de paso a la forma de Jordan, $J = P^{-1}AP$.

IV.35. Hallar la matriz de Jordan J de la siguiente matriz A, así como una matriz P de paso, $J = P^{-1}AP$:

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 2 & 0 & 2 \\ -1 & 1 & 0 \end{bmatrix}$$

IV.36. Hallar la matriz de Jordan, J, de la siguiente matriz A, así como una matriz P de paso de A a J, esto es. tal que $J = P^{-1}AP$:

$$A = \begin{bmatrix} -6 & 0 & 1 \\ -18 & -3 & 6 \\ -9 & 0 & 0 \end{bmatrix}$$

IV.37. Hallar la matriz de Jordan J, de la siguiente matriz A, así como una matriz P de paso de A a J, esto es, tal que $J = P^{-1}AP$:

$$A = \begin{bmatrix} 1 & 3 & 3 & -3 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 2 & 5 & 3 & -4 \end{bmatrix}$$

IV.38. Hallar la matriz de Jordan, J, de la siguiente matriz A, así como una matriz P de paso de A a J, esto es, tal que $J = P^{-1}AP$:

$$A = \begin{bmatrix} 1 & 4 & 0 & -6 & -8 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & -4 & 0 & 5 & 8 \\ 0 & 2 & 0 & -2 & -3 \\ 0 & -1 & 0 & 1 & 1 \end{bmatrix}$$

IV.39. Sea J una matriz de Jordan y J^t su matriz traspuesta compruébese que J y J^t son semejantes y hállense matrices de paso P, que permiten poner $J = P^{-1}J^tP$. (Indicación: sean $J_1, J_2, ..., J_q$ las cajas de Jordan de J; búsquese P entre las matrices diagonales por bloques $P_1, P_2, ..., P_q$ cuyos tamaños sean, respectivamente, iguales a los de $J_1, J_2, ..., J_q$. Póngase $PJ \Rightarrow J^tP$.)

IV.40. Sea J la matriz de Jordan de una matriz cuadrada A y sea P una matriz de paso de A a J, o sea, tal que $J = P^{-1}AP$. Hallar la matriz de Jordan J^* y una matriz de paso P^* de la matriz A^t, traspuesta de A. (Indicación: recúrrase al ejercicio anterior, utilizando una matriz P_0 que permita poner $J^t = P_0^{-1}JP_0$.)

IV.41. Sea A una matriz cuadrada de tamaño $n \times n$. Pruébese que si A tiene n autovalores (iguales o distintos) y tiene inversa A^{-1}, entonces ésta tiene también n autovalores, que son los inversos de los autovalores de A.

IV.42. Sea V el espacio vectorial real de los polinomios de grado menor o igual que n. Sea $f: V \to V$ el endomorfismo que a cada polinomio le atribuye su polinomio derivado. Hállese la matriz de Jordan J de f y una base B de V en la que J es la matriz de f.

IV.43. Hallar la forma canónica de Jordan, J, de la siguiente matriz A, así como una matriz P de paso de A a J, esto es, tal que, $J = P^{-1}AP$:

$$A = \begin{bmatrix} \lambda & n & & & O \\ & \lambda & n-1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ O & & & & \lambda \end{bmatrix}$$

IV.44. Sea A una matriz cuadrada, de tamaño $n \times n$, que es semejante a la siguiente matriz de Jordan J, esto es, existe una matriz regular P tal que $J = P^{-1}AP$. Hallar, en función de P, todas las matrices regulares Q tales que $J = Q^{-1}AQ$:

$$J = \begin{bmatrix} \lambda & 1 & & & O \\ & \lambda & 0 & & \\ & & \lambda_3 & & \\ & O & & \ddots & 0 \\ & & & & \lambda_n \end{bmatrix} \quad \lambda \neq \lambda_i \quad (i = 3, ..., n)$$

IV.45. Dado un polinomio, $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_kx^k$, en la indeterminada x, y dada una matriz cuadrada A, se define $p(A)$ como

$$p(A) = a_0I + a_1A + a_2A^2 + \cdots + a_kA^k$$

a) Si $p(x)$ y $q(x)$ son dos polinomios, obtener $(p+q)A$ y $(\alpha p)(A)$, donde α es un escalar, en función de $p(A)$ y $q(A)$. Relacionar $p(A)q(A)$ con $q(A)p(A)$.

b) Si A y A' son matrices cuadradas semejantes, con $A' = P^{-1}AP$, comprobar que A^n y A'^n también son semejantes y hallar la matriz de paso de A^n a A'^n. Idem para $p(A)$ y $p(A')$, siendo $p(x)$ un polinomio cualquiera.

c) Si A es una matriz cuadrada diagonal por bloques y estos bloques son las matrices cuadradas $A_1, A_2, ..., A_p$, compruébese que A^n y que $p(A)$ son matrices diagonales por bloques, hallando estos bloques, donde $p(x)$ es un polinomio.

IV.46. Teorema de Cayley-Hamilton. Sea A una matriz cuadrada compleja(*) y sea $p(\lambda)$ el polinomio característico de λ. Demuéstrese que $p(A) = O$ (toda matriz anula a su polinomio característico). Indicaciones: si λ_i $(i = 1, ..., p)$ son los autovalores de A, con multiplicidades m_i, póngase $p(x) = (-1)^n\pi(\lambda - \lambda_i)^{m_i}$; recúrrase a la forma de Jordan J de A; hágase uso de los apartados *b*) y *c*) del ejercicio anterior.

IV.47. Si A es una matriz regular, recurriendo al teorema de Cayley-Hamilton (ver el ejercicio anterior) y al polinomio característico de A, obtener una expresión para A^{-1}.

(*) El teorema es cierto también para matrices cuyos elementos pertenezcan a un cuerpo K cualquiera. Este caso general se puede reducir al caso complejo que aquí consideramos.

IV.48. Sea J una caja o bloque de Jordan de tamaño $n \times n$. Sea $p(x)$ un polinomio de grado h, que de acuerdo con la fórmula de Taylor se puede poner

$$p(x) = p(\lambda) + (x-\lambda)\frac{p'(\lambda)}{1!} + (x-\lambda)^2\frac{p''(\lambda)}{2!} + \\ + \cdots + (x-\lambda)^h\frac{p^{(h)}(\lambda)}{h!}$$

En función de $p(\lambda), p'(\lambda), p''(\lambda), \ldots, p^{(h)}(\lambda)$, hallar los elementos de la matriz $p(J)$ (véanse los ejercicios anteriores).

IV.49. Hallar la matriz de Jordan J de la siguiente matriz compleja A, así como una matriz P de paso de A a J, esto es, tal que $J = P^{-1}AP$:

$$A = \begin{bmatrix} 1+i & -1-i & 1 \\ i & i & 1 \\ -i & -1+i & 0 \end{bmatrix}$$

IV.50. Hallar la matriz de Jordan J de la siguiente matriz compleja A, así como una matriz P de paso de A a J, esto es, tal que $J = P^{-1}AP$:

$$A = \begin{bmatrix} 1+2i & 0 & 1-i & i \\ 0 & 1-i & 0 & 0 \\ -1 & 0 & i & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

IV.51. Sea J una matriz de Jordan. Descomponer J en suma de dos matrices, $J = D + N$, una diagonal y otra nilpotente que conmuten, $DN = ND$. Recurriendo al resultado anterior, pruébese que: si A es una matriz cuadrada de tamaño $n \times n$, cuyos elementos pertenecen a un cuerpo K, que tiene n autovalores de K (iguales o distintos), entonces A puede expresarse como suma $A = A_D + A_N$, donde A_D es diagonalizable y A_N es nilpotente, de manera que $A_D A_N = A_N A_D$.

IV.52. Sea A una matriz cuadrada idempotente, esto es, tal que $A^2 = A$. Recurriendo a la forma de Jordan J de A, comprobar que A es diagonalizable y obtener los valores posibles de los elementos de su forma diagonal.

IV.53. Sea A una matriz cuadrada tal que $A^2 = I$. Recurriendo a su forma de Jordan J, comprobar que A es diagonalizable y obtener los valores posibles de los elementos de su forma diagonal.

SOLUCIONES

IV.1. $\lambda = 1$ (con multiplicidad p) y $\lambda = 0$ (con multiplicidad q).

IV.2. $f(\bar{x}) = \lambda\bar{x} \Rightarrow (f \circ f)(\bar{x}) = \lambda^2\bar{x}$;
$f(\bar{x}) = \lambda\bar{x} \Rightarrow \bar{x} = \lambda f^{-1}(\bar{x})$

IV.3. *a)*
$$\begin{aligned}\det(AB - \lambda I) &= \det(AB - \lambda B^{-1}B) = \\ &= [\det(A - \lambda B^{-1})]\det B = \\ &= \det B\,[\det(A - \lambda B^{-1}) = \\ &= \det(BA - \lambda BB^{-1}) = \\ &= \det(BA - \lambda I)\end{aligned}$$

b) $\det(AB + \varepsilon B - \lambda I) = \det(BA + \varepsilon B - \lambda I)$, al tender ε a cero, como los coeficientes de estos polinomios varían continuamente con ε (son funciones continuas de ε), se obtiene que $\det(AB - \lambda I) = \det(BA - \lambda I)$. Nótese que si $\lambda \neq 0$ es el autovalor de B más próximo a 0, entonces B' es regular para $0 < \varepsilon < |\lambda|$.

IV.4. El polinomio característico $p(\lambda)$ de A es tal que $p(0) = \det(A) < 0$ y $p(\lambda) \to +\infty$ para $\lambda \to -\infty$; por tanto, $p(\lambda)$ tienen, al menos, una raíz positiva y otra negativa.

IV.5. $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 2$
$V_{\lambda=0} = \mathcal{V}(-3, 3, 2)$, $V_{\lambda=1} = \mathcal{V}(2, 0, -1)$,
$V_{\lambda=2} = \mathcal{V}(1, 1, 0)$,

f es diagonalizable.

IV.6. $\lambda = 2$ doble y $\lambda = -1$ simple
$V_{\lambda=2} = \mathcal{V}[(1, 0, 1), (1, 1, 0)]$,
$V_{\lambda=-1} = \mathcal{V}(1, -1, 1)$,

f es diagonalizable.

IV.7. $\lambda_1 = 2$ doble y $\lambda_2 = 1$ doble
$V_{\lambda=2} = \mathcal{V}(\bar{e}_1 - \bar{e}_2 + \bar{e}_3 + 2\bar{e}_4)$,
$V_{\lambda=1} = \mathcal{V}(\bar{e}_3, \bar{e}_1 - \bar{e}_2 + \bar{e}_3 + \bar{e}_4)$,

f no es diagonalizable.

IV.8. $\lambda_1 = a$ con multiplicidad $n - 1$
$\lambda_2 = a + n$ (simple)

$$\overrightarrow{V_{\lambda=a}}: \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ tales que } x_1 + x_2 + \cdots + x_n = 0$$

$$V_{\lambda=a+n}: \begin{bmatrix} x \\ x \\ \vdots \\ x \end{bmatrix} \quad (x \in \mathbb{R})$$

A es diagonalizable por semejanza.

IV.9. *a)* $\lambda_1 = 1$, $\lambda_2 = \lambda_3$, $\lambda_1 + \lambda_2 + \lambda_3 = 5$, luego $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = 2$.

b) $V_{\lambda_1} = \mathcal{V}(\bar{u})$, con $\bar{u}(6, 2, 5)$; $V_{\lambda_2} = \mathcal{V}(\bar{v}, \bar{w})$, con $\bar{v}(7, 0, 2)$ y $\bar{w}(11, -2, 0)$. En la base $(\bar{u}, \bar{v}, \bar{w})$ la matriz es diagonal, cambiando de base, se tiene:

$$A = \begin{bmatrix} 14 & 66 & -42 \\ 4 & 24 & -14 \\ 10 & 55 & -33 \end{bmatrix}$$

IV.10. Sea $P = I + E_{ij}$, donde E_{ij} tiene todos sus elementos nulos salvo el de lugar i, j que vale 1 (para $i \neq j$); de $PA = AP$ se deduce que el elemento de lugar i, j de A es nulo (con $i \neq j$). Las matrices A pedidas son las matrices escalares (tienen ceros fuera de su diagonal).

IV.11. A y A' son semejantes; P es la matriz que resulta de permutar las columnas i-ésima y j-ésima de la matriz unidad.

IV.12. $Q = RP$, donde R es una matriz regular que conmuta con A.

IV.13. $\det(A - \lambda I) = \lambda^n - 1$, que tiene n raíces distintas en $\mathbb{C}$, luego A es diagonalizable en $\mathbb{C}$. La forma diagonal es

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & O \\ & & \ddots & \\ O & & & \lambda_n \end{bmatrix}$$

$$\lambda_j = \cos\frac{2j\pi}{n} + i\,\text{sen}\,\frac{2j\pi}{n} \quad (\text{para } j = 1, 2, \ldots, n)$$

(i es la unidad imaginaria.)

IV.14. Si $A^k = O$ y λ es autovalor de A, existe $X \neq O$ tal que $AX = \lambda X$, luego $A^k X = \lambda^k X$, luego $\lambda = 0$.

Si los n autovalores de A son nulos, recurriendo a la forma triangular T de A (es $A = P^{-1}TP$ con T triangular; la diagonal de T está ocupada por los autovalores de A), como la diagonal de T tiene todos sus elementos nulos, es $T^n = O$, luego $A^n = O$.

IV.15. De ser $\alpha_1\bar{e}_1 + \alpha_2\bar{e}_2 + \cdots + \alpha_n\bar{e}_n = \bar{o}$, sería ($\lambda_i$ = = autovalor de f):

$$\alpha_1 + \alpha_2\lambda_i + \alpha_3\lambda_i^2 + \cdots + \alpha_n\lambda_i^{n-1} = 0$$

y este polinomio de grado $n-1$ tendría n raíces (los autovalores), luego sus coeficientes serían nulos.

IV.16. *a)* $\det(A - \lambda I) = [(4 - \alpha^2) - \lambda][(2\alpha + 4) - \lambda] \cdot [(4 - \alpha] - \lambda]$

$$\lambda_1 = 4 - \alpha^2, \quad \lambda_2 = 2\alpha + 4, \quad \lambda_3 = 4 - \alpha$$

Para $\alpha = 0$, f no es diagonalizable.
Para $\alpha \neq 0$, f es diagonalizable.

b) $\alpha = 1 \rightarrow D = P^{-1}AP$ con

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix} \text{ y } P = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$\alpha = 2 \rightarrow D = P^{-1}AP$ con

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 2 \end{bmatrix} \text{ y } P = \begin{bmatrix} 1 & 1 & 6 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

IV.17. *a)* $\lambda_1 = -1$ (simple) y $\lambda_2 = 1$ doble.

b) f diagonalizable $\Leftrightarrow \alpha = 0$

Si $\alpha = 0$,

$$V_{\lambda=-1} = \mathcal{V}(\bar{e}_2 - \bar{e}_3)$$

y

$$V_{\lambda=1} = \mathcal{V}(\bar{e}_1 + 2\bar{e}_2, \bar{e}_1 + 2\bar{e}_3)$$

Si $\alpha \neq 0$,

$$V_{\lambda=-1} = \mathcal{V}(\bar{e}_2 - \bar{e}_3)$$

y

$$V_{\lambda=1} = \mathcal{V}(2\bar{e}_1 + 3\bar{e}_2 + \bar{e}_3)$$

c) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ base: $\bar{u}_1 = \bar{e}_2 - \bar{e}_3$, $\bar{u}_2 = \bar{e}_1 + 2\bar{e}_2$, $\bar{u}_3 = \bar{e}_1 + 2\bar{e}_3$

IV.18. *a)* $f(1, 0, 1) = (-1, 0, -1)$; $\lambda = -1$ doble

$$A = \begin{bmatrix} t & -2t & -1-t \\ 0 & -1 & 0 \\ -1-t & 2t & t \end{bmatrix} \quad (t \neq -1)$$

b) $t = 0$.

IV.19. Por ser $A^t = -A$, es $X^tAX = 0$ para toda columna X, pues:

$$X^tAX = (X^tAX)^t = X^tA^tX = -X^tAX$$

Sea $\lambda = \alpha + \beta i$ un autovalor de A y $X + iY$ un vector propio correspondiente a λ, es

$$A(X + iY) = \lambda(X + iY)$$

luego

$$\left.\begin{aligned} AX &= \alpha X - \beta Y \\ AY &= \alpha Y + \beta X \end{aligned}\right\}$$

de donde:

$$X^tAX = \alpha X^tX - \beta X^tY$$
$$Y^tAY = \alpha Y^tY + \beta Y^tX$$

Sumando, es $O = \alpha(X^tY + Y^tY)$, luego $\alpha = 0$ y $\lambda = \beta i$.

IV.20 *a)* $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}$ $\lambda_1 = 0$, $\lambda_2 = 6$, $\lambda_3 = 3$

$$V_{\lambda=0} = \mathcal{V}(2\bar{e}_1 - 2\bar{e}_2 - \bar{e}_3)$$
$$V_{\lambda=3} = \mathcal{V}(\bar{e}_1 + 2\bar{e}_2 - 2\bar{e}_3)$$
$$V_{\lambda=6} = \mathcal{V}(2\bar{e}_1 + \bar{e}_2 + 2\bar{e}_3)$$

b) $D = P^{-1}AP$ con

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad P = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ -1 & -2 & 2 \end{bmatrix}$$

nueva base:

$$\bar{u}_1 = \frac{1}{3}(2\bar{e}_1 - 2\bar{e}_2 - \bar{e}_3)$$

$$\bar{u}_2 = \frac{1}{3}(\bar{e}_1 + 2\bar{e}_2 - 2\bar{e}_3)$$

$$\bar{u}_3 = \frac{1}{3}(2\bar{e}_1 + \bar{e}_2 + 2\bar{e}_3)$$

IV.21. La matriz de f es

$$A = \frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$$

(que es simétrica y ortogonal).

a) f es ortogonalmente diagonalizable (A es simétrica).
b) f es una transformación ortogonal (A es ortogonal).
c) $\lambda_1 = 1$ (simple) y $\lambda_2 = -1$ (doble)

$$V_{\lambda=1} = \mathcal{V}(1, 1, 1)$$
$$V_{\lambda=-1} = \mathcal{V}[(1, -1, 0), (1, 0, -1)],$$

d)
$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

base:

$$\bar{u}_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$$

$$\bar{u}_2 = \frac{1}{\sqrt{2}}(0, 1, -1)$$

$$\bar{u}_3 = \frac{1}{\sqrt{2}}(-2, 1, 1)$$

e) f es la simetría (ortogonal) respecto de la recta que determina el vector (1, 1, 1).

IV.22. $\lambda_1 = \alpha$, $\lambda_2 = \alpha \pm \sqrt{\beta\gamma}$
a) Hay autovalor triple si $\beta\gamma = 0$.
b) $\lambda_1 = \alpha$, $\lambda_2 = \alpha + h$, $\lambda_3 = \alpha - h$

$$V_{\lambda 1} = \mathcal{V}(-\beta\bar{e}_1 + \gamma\bar{e}_3)$$
$$V_{\lambda 2} = \mathcal{V}(\beta\bar{e}_1 + h\bar{e}_2 + \gamma\bar{e}_3)$$
$$V_{\lambda=3} = \mathcal{V}(\beta\bar{e}_1 - h\bar{e}_2 + \gamma\bar{e}_3)$$

c) $(-\beta\bar{e}_1 + \gamma\bar{e}_3,\ \beta\bar{e}_1 + h\bar{e}_2 + \gamma\bar{e}_3,\ \beta\bar{e}_1 - h\bar{e}_2 + \gamma\bar{e}_3)$ ha de ser un sistema ortogonal, o sea si $\beta^2 = \gamma^2 + \beta^2 + \gamma^2 = h^2$, o sea, para $\beta = \gamma$.

IV.23. *a*) $\alpha = -2$, $\beta = -2$, $\gamma = -3$
b) $\lambda = 1$ (simple), $\lambda = -1$ doble

$$V_{\lambda=1} = \mathcal{V}(1, 1, 1)$$
$$V_{\lambda=-1} = \mathcal{V}[(1, 0, 1), (0, 1, 1)]$$

c)
$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

d) No es ortogonalmente diagonalizable, pues $V_{\lambda=1}$ y $V_{\lambda=-1}$ no son ortogonales.

IV.24.
$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}$$

a) f es ortogonalmente diagonalizable (A es simétrica); f no es una transformación ortogonal (A no es ortogonal).
b) $\lambda_1 = 0$, $\lambda_2 = 3$, $\lambda_3 = 6$

$$V_{\lambda_1} = \mathcal{V}(2, -2, -1)$$
$$V_{\lambda_2} = \mathcal{V}(1, 2, -2)$$
$$V_{\lambda_3} = \mathcal{V}(2, 1, 2)$$

c)
$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

base:

$$\bar{u}_1 = \frac{1}{3}(2, -2, -1)$$

$$\bar{u}_2 = \frac{1}{3}(1, 2, -2)$$

$$\bar{u}_3 = \frac{1}{3}(2, 1, 2)$$

IV.25. *a*) $\lambda = 2$ triple y $\lambda_2 = -2$ (simple).
b) $\alpha = 0$, único valor para el que dim $V_{\lambda=2} = 3$.
c) Es ortogonalmente diagonalizable para $\alpha = 0$

$$D=\begin{bmatrix} 2 & & & \\ & 2 & & O \\ & & 2 & \\ & O & & -2 \end{bmatrix}$$

$$P=\begin{bmatrix} 1/\sqrt{2} & 0 & 1/2 & -1/2 \\ 1/\sqrt{2} & 0 & -1/2 & 1/2 \\ 0 & 1/\sqrt{2} & 1/2 & 1/2 \\ 0 & -1/\sqrt{2} & 1/2 & 1/2 \end{bmatrix}$$

d) f es la simetría respecto del plano ortogonal al vector $(-1, 1, 1, 1)$.

e) $G = A(0) + hI$ es simétrica; debe ser definida positiva, esto es, ha de tener autovalores positivos; sus autovalores son $2+h$ (triple) y $-2+h$ (simple); por tanto, $h = 3$.

IV.26. Los autovalores de la matriz de ω (en base canónica) son $\lambda_1 = 1$, $\lambda_2 = \sqrt{3}$ y $\lambda_3 = -\sqrt{3}$. Los subespacios propios son:

$$V_{\lambda_1} = \mathcal{V}(1, 0, -1)$$
$$V_{\lambda_2} = \mathcal{V}(1, -1-\sqrt{3}, 1)$$
$$V_{\lambda_3} = \mathcal{V}(1, -1+\sqrt{3}, 1)$$

$$\omega(x_1, x_2, x_3) = x_1^2 + \sqrt{3}x_2^2 - \sqrt{3}x_3^2$$

donde x_1, x_2, x_3 son las coordenadas en la base ortonormal $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$:

$$\bar{u}_1 = \frac{1}{\sqrt{2}}(1, 0, -1)$$

$$\bar{u}_2 = \frac{1}{\sqrt{7+2\sqrt{3}}}(1, -1-\sqrt{3}, 1)$$

$$\bar{u}_3 = = \frac{1}{\sqrt{7-2\sqrt{3}}}(1, -1+\sqrt{3}, 1)$$

IV.27 Autovectores y subespacios propios de A

$$\lambda_1 = 3, \quad \lambda_2 = 6, \quad \lambda_3 = -2$$

$$V_{\lambda_1} = \mathcal{V}(1, -1, 1)$$
$$V_{\lambda_2} = \mathcal{V}(1, -1, -2)$$
$$V_{\lambda_3} = \mathcal{V}(1, 1, 0)$$

$$\bar{u}_1 = \frac{1}{\sqrt{3}}(1, -1, 1)$$

$$D=\begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -2 \end{bmatrix} \quad \bar{u}_2 = \frac{1}{\sqrt{6}}(1, -1, -2)$$

$$\bar{u}_3 = \frac{1}{\sqrt{2}}(1, 1, 0)$$

IV.28. $A'(h)$ es definida positiva, definida negativo y no definida según que, respectivamente, sea $h < \alpha$, $h > \beta$ y $\alpha \leqslant h \leqslant \beta$.

IV.29. Autovalores de A: $\lambda_1 = 1$ triple y $\lambda_2 = 5$ (simple)

$$V_{\lambda_1} = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 / x_1 + x_2 + x_3 + x_4 = 0\}$$

$$V_{\lambda_2} = \mathcal{V}(1, 1, 1, 1)$$

$$D=\begin{bmatrix} 1 & & & \\ & 1 & & O \\ & & 1 & \\ & O & & 5 \end{bmatrix}$$

$$P=\begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

La matriz A es definida positiva, pues sus autovalores son positivos.

IV.30. Sean λ_i los autovalores de A; sea (X_i) una base ortogonal de autovectores de A; toda columna X es $X = \Sigma \alpha_i X_i$

$$(AX)X^t = (\Sigma \alpha_i AX_i)(\Sigma \alpha_j X_j^t = (\Sigma \alpha_i \lambda_i X_i)(\Sigma \alpha_j X_j^t) =$$

$$= \Sigma \alpha_i^2 \lambda_i X_i X_i^t$$

$$a_1 XX^t = \Sigma a_i \alpha_i^2 X_i X_i^t \leqslant \Sigma \alpha_i^2 \lambda_i X_i X_i^t = AXX^t$$

Análogamente $a_2XX^t \geqslant AXX^t$. Para toda columna X es

$$\left.\begin{array}{l} a_1XX^t \leqslant AXX^t \leqslant a_2XX^t \\ b_1XX^t \leqslant BXX^t \leqslant b_2XX^t \end{array}\right\} \quad [1]$$

$$h_1XX^t \leqslant (A+B)XX^t \leqslant h_2XX^t \quad [2]$$

(se han llamado h_1 y h_2 al menor y al mayor de los autovalores de $A+B$). Sumando las relaciones [1], se obtiene que

$$(a_1+b_1)XX^t \leqslant (A+B)XX^t \leqslant (a_2+b_2)XX^t \quad [3]$$

Como existen ciertas columnas X_1 y X_2 (vectores propios de $A+B$ asociados a h_1 y h_2) tales que

$$h_1X_1X_1^t = (A+B)X_1X_1^t$$

y

$$h_2X_2X_2^t = (A+B)X_2X_2^t \quad [4]$$

de [2], [3] y [4] resulta que

$$[h_1, h_2] \subset [a_1+b_1, a_2+b_2]$$

IV.31. El endomorfismo f diagonalizable. En tal caso, el producto escalar es aquel que hace ortonormal a una base en lo que f tenga matriz diagonal.

IV.32.

$$\begin{bmatrix} 3 & 1 & & & \\ & 3 & 1 & & \\ & & 3 & & \\ & & & 3 & \\ & & & & 2 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & 2 & & & \\ & & 3 & 1 & \\ & & & 3 & \\ & & & & 2 \end{bmatrix}$$

IV.33. $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -4 & 1 & 0 \\ 0 & -4 & 1 \\ 0 & 0 & -4 \end{bmatrix}$

IV.34. No depende de β.

Si $\alpha = 0$,

$$J = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Si $\alpha \neq 0$,

$$J = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{y} \quad P = \begin{bmatrix} \alpha & 0 & \alpha\beta \\ 0 & 1 & 3\beta \\ 0 & 0 & 9 \end{bmatrix}$$

IV.35. $J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix} \quad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{bmatrix}$

IV.36. $J = \begin{bmatrix} -3 & 1 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix} \quad P = \begin{bmatrix} 1 & 0 & 0 \\ 6 & 2 & 1 \\ 3 & 1 & 0 \end{bmatrix}$

IV.37.

$$J = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

$$P = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

IV.38.

$$J = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P = \begin{bmatrix} 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \\ 1 & -3 & 0 & -4 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix}$$

IV.39.

$$P_i = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & 0 & \cdots & a_1 & a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & a_1 & \cdots & a_{h-3} & a_{h-2} \\ 0 & a_1 & a_2 & \cdots & a_{h-2} & a_{h-1} \\ a_1 & a_2 & a_3 & \cdots & a_{h-1} & a_h \end{bmatrix}$$

con $a_1 \neq 0$.

IV.40. $J=(P_0P^t)A^t(P_0P^t)^{-1}$, $J^*=J$ y $P^*=P_0P^t$.

IV.41. Recurrir a que, para cierta matriz P, la matriz $P^{-1}AP$ es triangular y su diagonal está ocupada por los autovalores de A.

IV.42.
$$J=\begin{bmatrix} 0 & 1 & & O \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & O & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$$

y

$$B=\left(1, \frac{x}{1!}, \frac{x^2}{2!}, \ldots, \frac{x^n}{n!}\right)$$

IV.43.
$$J=\begin{bmatrix} \lambda & 1 & & & O \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & O & & & \lambda \end{bmatrix}$$

$$P=\begin{bmatrix} n! & & & O \\ & (n-1)! & & \\ & & \ddots & \\ & O & & 1! \\ & & & & 0! \end{bmatrix}$$

IV.44.
$$Q=P\begin{bmatrix} a_1 & b & & O \\ & a_2 & & \\ & & \ddots & \\ & O & & a_n \end{bmatrix} \quad \text{con} \quad a_1=a_2$$

IV.45. *a*) $(p+q)(A)=p(A)+q(A)$, $(pq)(A)=p(A)q(A)$, $(\alpha p)(A)=\alpha p(A)$, $p(A)q(A)=q(A)p(A)$

b) Por inducción, $A'^n=P^{-1}A^nP$. Aplicando lo anterior, $p(A')=P^{-1}p(A)P$.

c) Bloques diagonales de A^n:

$$A_1^n, A_2^n, \ldots, A_p^n$$

Bloques diagonales de $p(A)$:

$$p(A_1), p(A_2), \ldots, p(A_n)$$

IV.46. $p(A)=P^{-1}p(J)P$;

$$p(J)=(-1)^n(J-\lambda_1 I)^{m_1}\ldots(J-\lambda_p I)^{m_p};$$

Si los bloques de Jordan de J son $J_1, \ldots, J_h$, con

$$J_i=\begin{bmatrix} \lambda_i & 1 & & O \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & O & & & \lambda_i \end{bmatrix} \qquad (J_i-\lambda_i I)^{m_i}=O$$

Los bloques de $p(J)$ son los $p(J_i)$, que son todos nulos por tener un factor nulo cada uno de ellos. Luego $p(J)=O$ y $p(A)=O$.

IV.47. Si el polinomio característico de A es el

$$p(\lambda)=a_0+a_0+a_1\lambda+a_2\lambda^2+\cdots+a_{n-1}\lambda^{n-1}+(-1)^n\lambda^n$$

es $P(A)=O$; multiplicando por A^{-1}, se obtiene:

$$A^{-1}=-\frac{1}{a_0}(a_1I+a_2A+a_3A^2+\cdots+a_{n-1}A^{n-2}+\\+(-1)^nA^{n-1})$$

(nótese que $a_0=\det A\neq 0$).

IV.48. Como

$$p(J)=p(\lambda)I+\frac{p'(\lambda)}{1!}(J-\lambda I)+\cdots+\\+\frac{p^{(h)}(\lambda)}{h!}(J-\lambda I)^h$$

es:

$$p(J)=\begin{bmatrix} p(\lambda) & \frac{p'(\lambda)}{1!} & \frac{p''(\lambda)}{2!} & \cdots & \frac{p^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & p(\lambda) & \frac{p'(\lambda)}{1!} & \cdots & \frac{p^{(n-2)}(\lambda)}{(n-2)!} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & p(\lambda) \end{bmatrix}$$

IV.49.
$$J=\begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 1 \\ 0 & 0 & i \end{bmatrix} \quad \text{y} \quad P=\begin{bmatrix} 1 & i & -1 \\ 0 & 1 & 0 \\ i & 1 & 1+i \end{bmatrix}$$

IV.50.
$$J=\begin{bmatrix} 1+i & 0 & 0 & 0 \\ 0 & 1-i & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P=\begin{bmatrix} 1 & 0 & i & 0 \\ 0 & 1 & 0 & 0 \\ i & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

IV.51. D es la matriz diagonal que tiene la misma diagonal que J; $N=J-D$. Si J es la matriz de Jordan de A, con $J=P^{-1}AP$, tómese $A_D=PDP^{-1}$ y $A_N=PNP^{-1}$.

IV.52. $J^2=J$, luego J es diagonal y en su diagonal sólo hay elementos valiendo 1 ó 0.

IV.53. $J^2=I$, luego J es diagonal y en su diagonal sólo hay elementos valiendo 1 ó -1.

PARTE V

ES llegado el momento de cambiar de tercio y ocuparnos de la geometría. Sepa, empero, el lector poco avisado, que seguimos lidiando el mismo toro, pues 'álgebra lineal' y 'geometría cartesiana', a más de emparentar por todas partes, son disciplinas que congenian y, de antiguo, están bien avenidas.
Aunque a algunos les llegue a sorprender, pareciéndoles chocante y, quizá, exagerado, lo cierto es que, si echamos mano de los vectores, como haremos, resultará que, a la postre, vamos a precisar de un único elemento nuevo, de un punto, de uno solo, para poder montar, sobre él, nuestro tinglado. Tomando a este punto como origen, anclaje o asidero, en el que apoyar los vectores, y mirando atentamente a sus extremos, descubriremos que allí están ubicados los puntos, todos ellos, los que pueblan y llenan el espacio.
A quienes, cuando se enfrentaron con los vectores del álgebra, salieron con pie firme de la empresa, poco ha de costarles, pues, vérselas ahora con los puntos, con las rectas y con los planos que les saldrán al paso en geometría.

11. Los espacios geométricos E_2 y E_3.
12. Espacios puntuales (caso general). Ejercicios y problemas.

CAPITULO

11

Los espacios geométricos E_2 y E_3

Para iniciarnos debidamente en el estudio de la geometría, como en el de cualquier otra disciplina, es obligado conocer el terreno que se pisa. Aquí, este terreno es el «plano geométrico» o el «espacio geométrico» (según que la dimensión sea dos o tres), que con mayor propiedad se llaman «plano afín» y «espacio afín». Nuestro punto de partida ha de ser, pues, un conocimiento preciso de los conceptos de punto, recta, plano; sin ello, mal vamos a poder estudiar las relaciones y propiedades de estos elementos. Al lector se le sugieren dos modos de abordar estas cuestiones:

En uno de ellos, el más pedagógico, se comenzará con una presentación esquemática, lo más extractada posible, de la fundamentación axiomática de los espacios geométricos, que permita, no obstante, estudiar de modo satisfactorio los problemas afines y métricos usuales de la geometría ordinaria (bi y tridimensional). Más adelante y tras un análisis sosegado de los axiomas, se podrán estudiar los espacios afines y afines euclídeos generales. Quienes deseen utilizar este procedimiento, podrán seguir, sin alteraciones, el orden de las páginas que siguen.

También se puede empezar con un estudio detallado de la axiomática de los espacios afines, que permita abordar los problemas geométricos, afines y euclídeos, tanto en los espacios E_2 y E_3 como en espacios afines cualesquiera. Para utilizar esta vía, comiéncese con el primer apartado (el α) del próximo capítulo (el 12); después se puede, según se desee, o continuar con aquel capítulo y luego abordar el estudio de éste, si se quiere ir de lo general a lo particular, o invertir el orden, si se estima que lo mejor es comenzar con las cosas más concretas para terminar con aquello que tiene mayor grado de abstracción.

α AXIOMAS Y DEFINICIONES

A estas alturas, no podemos asentar el estudio de la Geometría en los conceptos intuitivos de punto, recta y plano que se utilizan, atinadamente, en la enseñanza media. Procediendo como se suele hacer en la matemática, fundamentaremos todo en unos axiomas que recojan las propiedades que se les suponen a los puntos de la Geometría elemental. Los axiomas, que no han «caído del cielo», que son el fruto de un proceso de abstracción, que se verifican en particular para el modelo intuitivo del que salieron, que permiten generalizar con acierto y provecho la Geometría elemental, a la postre se aceptan, se admiten sin demostración, son proposiciones ciertas por definición.

11.1. LOS ESPACIOS GEOMETRICOS BI Y TRIDIMENSIONALES

Para definir estos espacios geométricos, formados por puntos, vamos a recurrir a los correspondientes axiomas. Estos se han establecido de modo que recojan las propiedades que, en Geometría elemental, gobiernan las relaciones entre puntos y vectores. De este modo, podremos recurrir a cuanto ya sabemos acerca de los vectores. Así, pues, «trasladaremos» a la geometría las propiedades y los resultados del álgebra de vectores.

PLANO AFIN (E_2) Y ESPACIO AFIN (E_3)

[198]

El *espacio afín* (tridimensional) está constituido por los siguientes elementos. Un conjunto E_3 (a cuyos elementos se les llama *puntos*), el espacio vectorial $\mathbb{R}^3$ y una aplicación, que a cada par de puntos (P, Q) le asigna un vector $\bar{v} \in \mathbb{R}^3$, que se denota poniendo

$$\bar{v} = \overrightarrow{PQ} \quad \text{o también} \quad Q = P + \bar{v} \qquad [1]$$

todo ello de manera que se verifiquen las dos condiciones siguientes: 1.ª) para cada $P \in E_3$ y cada $\bar{v} \in \mathbb{R}^3$, existe un y sólo un $Q \in E_3$ que satisface a [1]; y 2.ª) dados tres puntos cualesquiera $P, Q, R \in E_3$, se verifica que

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR} \quad \text{(relación de Chasles)}$$

(esta última relación puede sustituirse por la $(P + \bar{u}) + \bar{v} = P + (\bar{u} + \bar{v})$, para cualesquiera $P \in E_3$ y $\bar{u}, \bar{v} \in \mathbb{R}^3$).

Un conjunto E_3 es el *plano afín* (espacio de dimensión dos) si para él se verifica lo anterior sustituyendo $\mathbb{R}^3$ por $\mathbb{R}^2$ (y E_3 por E_2).

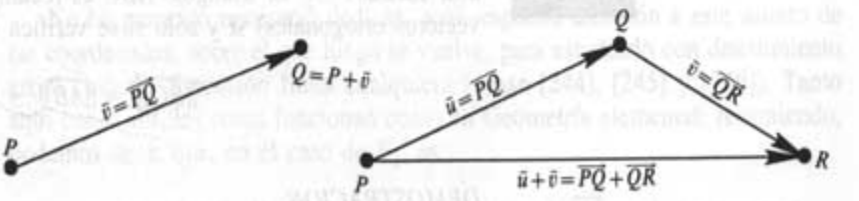


Figura 11.1.

Más adelante (véase [237] y [238]) se insistirá sobre esta definición, generalizándola, matizándola y completándola. Allí se puede acudir, ya ahora, si se

desea profundizar en ella y ver, así, cómo con esta definición se mantienen las propiedades que conocemos de geometría elemental. Así, por ejemplo, en [239] puede verse que: para cualesquiera que sean los puntos P, Q, P' y Q' (de E_2 o de E_3), se verifica que:

$$\overrightarrow{PQ} = \vec{0} \Leftrightarrow P = Q; \quad \overrightarrow{PQ} = -\overrightarrow{QP}; \quad \overrightarrow{PQ} = \overrightarrow{P'Q'} \Rightarrow \overrightarrow{PP'} = \overrightarrow{QQ'}$$

EL PLANO (E_2) Y EL ESPACIO (E_3) AFIN-EUCLIDEOS. DISTANCIA

[199]

El espacio afín E_3 (o el plano afín E_2) se dice que es *euclídeo* si se dispone del producto escalar entre sus vectores libres, es decir, si al espacio vectorial $\mathbb{R}^3$ (o $\mathbb{R}^2$) se le considera con su estructura euclídea.

Se llama *distancia* del punto P al punto Q, ambas de E_3 o de E_2 (euclídeos), al número real

$$d(P, Q) = \|\overrightarrow{PQ}\|$$

Cuando (en [240] y [241]) se estudien estas cuestiones con mayor generalidad, se comprobará que la distancia aquí definida goza de las propiedades que ya conocemos de geometría elemental. En concreto, podríamos ya comprobar (siguiendo los mismos pasos que en [241]) que, para cualesquiera que sean los puntos P, Q y P de E_2 o de E_3 (euclídeos), se verifica que:

- $d(P, Q) > 0$ si $P \neq Q$; $\quad d(P, P) = 0$
- $d(P, Q) = d(Q, P)$
- $d(P, R) \leq d(P, Q) + d(Q, R)$ (desigualdad triangular)

TEOREMA DE PITAGORAS

[200]

Sean A, B y C tres puntos distintos del espacio afín euclídeo E_3 o del plano afín euclídeo E_2. El triángulo ABC es rectángulo en A (esto es $\overrightarrow{AB}$ y $\overrightarrow{AC}$ son vectores ortogonales) si y sólo si se verifica que

$$\|\overrightarrow{BC}\|^2 = \|\overrightarrow{AB}\|^2 + \|\overrightarrow{AC}\|^2$$

DEMOSTRACION

Para cualesquiera que sean los puntos A, B y C, se verifica que

$$\overrightarrow{BC} = \overrightarrow{AC} - \overrightarrow{AB}$$

- Si la referencia es rectangular, entonces la distancia entre los puntos $X(x_i)$ e $Y(y_i)$ es $\sqrt{(y_1-x_1)^2+(y_2-x_2)^2+(y_3-x_3)^2}$.
- Sean (x_1, x_2, x_3) las coordenadas de un punto $X \in E_3$ en la referencia $(O; \bar{e}_1, \bar{e}_2, \bar{e}_3)$. Considérese otra referencia $(C; \bar{u}_1, \bar{u}_2, \bar{u}_3)$, donde se conocen $C(c_1, c_2, c_3)$ y $\bar{u}_j = q_{1j}\bar{e}_1 + q_{2j}\bar{e}_2 + q_{3j}\bar{e}_3$ (para $j = 1, 2, 3$). Si (x'_1, x'_2, x'_3) son las coordenadas de X en la nueva referencia, entonces

$$x_i = c_i + q_{i1}x'_1 + q_{i2}x'_2 + q_{i3}x'_3$$

(para $i = 1, 2, 3$). Este resultado se obtiene fácilmente (véase también [245]):

$$\sum_{i=1}^{3} x_i\bar{e}_i = \overrightarrow{OX} = \overrightarrow{OC} + \overrightarrow{CX} = \sum_{i=1}^{3} c_i\bar{e}_i + \sum_{j=1}^{3} x'_j\bar{u}_j = \sum_{i=1}^{3} c_i\bar{e}_i + \sum_{j=1}^{3} x'_j\left(\sum_{i=1}^{3} q_{ij}\bar{e}_i\right) =$$

$$= \sum_{i=1}^{3}\left(c_i + \sum_{j=1}^{3} x'_j q_{ij}\right)\bar{e}_i$$

luego

$$x_i = c_i + \sum_{j=1}^{3} q_{ij}x'_j \quad (\text{para } i = 1, 2, 3)$$

Si las dos referencias anteriores son rectangulares, entonces la matriz $Q = [q_{ij}]$ es ortogonal.

11.2. LAS RECTAS Y LOS PLANOS

Si V es un espacio vectorial de dimensión 1 ó 2, al «aplicar» los vectores de V en un punto fijo P (origen de todos ellos), los extremos de dichos vectores recorren una recta o un plano, respectivamente.

RECTAS (EN E_2 Y EN E_3)

[202]

Dados un punto P, del espacio E_3, y un vector no nulo $\bar{u} \in \mathbb{R}^3$, se llama *recta* que pasa por P y tiene la dirección de $\bar{u}$ al conjunto formado por los puntos $X \in E_3$ tales que $\overrightarrow{PX} = \lambda\bar{u}$ para $\lambda \in \mathbb{R}$, esto es, tales que:

$$X = P + \lambda\bar{u}, \quad \text{para } \lambda \in \mathbb{R}$$

(*ecuación vectorial* de la recta)

Por dos puntos distintos P, $Q \in E_3$ pasa una sola recta, que es la que forman los puntos $X = P + \lambda\overrightarrow{PQ}$ cuando $\lambda \in \mathbb{R}$.

De igual modo se definen las rectas de E_2: sustituyendo E_3 y $\mathbb{R}^3$ por E_2 y $\mathbb{R}^2$.

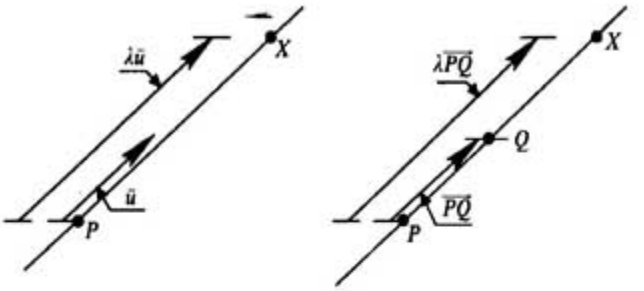


Figura 11.3.

OBSERVACIONES

- La recta que pasa por el punto P y tiene la dirección del vector $\bar{u}$, esto es, el conjunto $\{X = P + \lambda\bar{u} / \lambda \in \mathbb{R}\}$, se denota poniendo $P + \mathcal{V}(\bar{u})$, donde $\mathcal{V}(\bar{u}) = \{\lambda\bar{u} / \lambda \in \mathbb{R}\}$ es el subespacio que engendra $\bar{u}$. A este subespacio se le llama *dirección* de la recta.
- Según se habrá notado, para señalar que un punto P pertenece a una recta, se dice que ésta pasa por P.
- Según se comprueba con facilidad, la recta que pasa por P y tiene la dirección de $\bar{u}$ y la recta que pasa por Q y tiene la dirección de $\bar{v}$ son la misma si y sólo si $\bar{u}$, $\bar{v}$ y $\overrightarrow{PQ}$ son vectores proporcionales.
- Se llama *segmento* que tiene por origen y extremo a los puntos P y Q al conjunto $[P, Q] = \{X = P + \lambda\overrightarrow{PQ} / 0 \leqslant \lambda \leqslant 1\}$. Se llama *punto medio* de los P y Q al $M = P + \frac{1}{2}\overrightarrow{PQ}$.

EJERCICIO

Sean dados cuatro puntos A, B, C y D (de E_2 o de E_3) que forman un paralelogramo, esto es, tales que $\overrightarrow{AB} = \overrightarrow{DC}$. Compruébese que sus diagonales se cortan en su punto medio.

RESOLUCION

Nótese que, por ser $\overrightarrow{AB} = \overrightarrow{DC}$, también es $\overrightarrow{AD} = \overrightarrow{BC}$. Si M_1 es el punto medio de AC y M_2 es el punto medio de PB, hay que comprobar que $M_1 = M_2$. Así es, ya que

$$M_2 = D + \tfrac{1}{2}\overrightarrow{DB} = (A + \overrightarrow{AD}) + \tfrac{1}{2}(\overrightarrow{DA} + \overrightarrow{AB}) =$$
$$= A + \tfrac{1}{2}(\overrightarrow{AD} + \overrightarrow{AB}) = A + \tfrac{1}{2}\overrightarrow{AC} = M_1$$

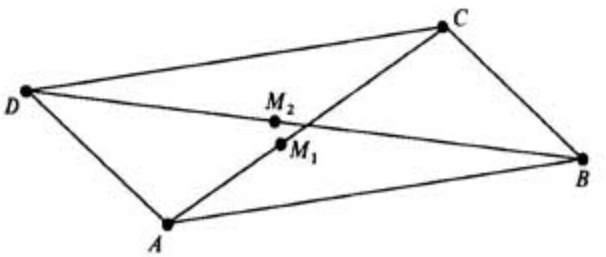


Figura 11.4.

PLANOS (DE E_3)

[203]

Dados un punto P, del espacio E_3, y dos vectores $\bar{u}, \bar{v} \in \mathbb{R}^3$ linealmente independientes, se llama *plano* que pasa por P y tiene a los $\bar{u}$ y $\bar{v}$ como vectores de dirección al conjunto formado por los puntos $X \in E_3$ tales que $\overrightarrow{PX} = \lambda\bar{u} + \mu\bar{v}$ para $\lambda, \mu \in \mathbb{R}$, esto es, tales que:

$$X = P + \lambda\bar{u} + \mu\bar{v}, \quad \text{para } \lambda, \mu \in \mathbb{R}$$

(*ecuación vectorial* del plano)

Por tres puntos no alineados $P, Q, R \in E_3$ pasa un solo plano, que es el que forman los puntos

$$X = P + \lambda\overrightarrow{PQ} + \mu\overrightarrow{PR}, \quad \text{para } \lambda, \mu \in \mathbb{R}$$

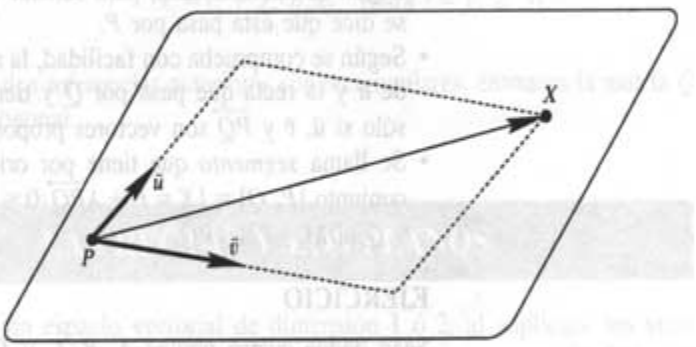


Figura 11.5.

OBSERVACIONES

- El plano que pasa por el punto P y tiene como vectores de dirección a los $\bar{u}$ y $\bar{v}$ (independientes), esto es, el conjunto $\{X = P + \lambda\bar{u} + \mu\bar{v}/\lambda, \mu \in \mathbb{R}\}$, se suele denotar poniendo $P + \mathcal{V}(\bar{u}, \bar{v})$, donde $\mathcal{V}(\bar{u}, \bar{v})$ es el subespacio vectorial (bidimensional) que engendran $\bar{u}$ y $\bar{v}$. A este subespacio se le llama *dirección* del plano.
- En lugar de decir que un punto P pertenece a un cierto plano, se acostumbra a decir que el plano pasa por P.
- Es fácil comprobar que dos planos $P_1 + \mathcal{V}(\bar{u}_1, \bar{v}_1)$ y $P_2 + \mathcal{V}(\bar{u}_2, \bar{v}_2)$ coinciden si y sólo si $\overrightarrow{P_1P_2} \in \mathcal{V}(\bar{u}_1, \bar{v}_1) = \mathcal{V}(\bar{u}_2, \bar{v}_2)$.

EJERCICIOS

1. Sea dada la recta $r: X = P + \lambda\bar{u}$ (pasa por P y tiene la dirección de $\bar{u}$), del espacio E_3. Hallar dos planos que se corten según r.

2. Dados dos planos que pasan, ambos, por un mismo punto P, compruébese que, si no son el mismo, se cortan según una recta.

RESOLUCION

1. Sean $\bar{v}_1$ y $\bar{v}_2$ dos de los muchos vectores tales que $(\bar{u}, \bar{v}_1, \bar{v}_2)$ es una base de $\mathbb{R}^3$. Los planos $P + \mathcal{V}(\bar{u}, \bar{v}_1)$ y $P + \mathcal{V}(\bar{u}, \bar{v}_2)$, esto es los que tienen por ecuaciones (para $\lambda, \mu_1, \mu_2 \in \mathbb{R}$) a:

$$X = P + \lambda\bar{u} + \mu_1\bar{v}_1 \quad \text{y} \quad X = P + \lambda\bar{u} + \mu_2\bar{v}_2$$

se cortan obviamente a lo largo de la recta r dada.

2. Los planos dados se pueden poner en la forma $P + U_1$ y $P + U_2$, donde U_1 y U_2 son sus direcciones (subespacios vectoriales de dimensión 2). Los planos son distintos si $U_1 \neq U_2$ y, entonces $U = U_1 \cap U_2$ es un subespacio de dimensión 1. Como la intersección de los planos es el conjunto $P + U$, este es, pues, una recta.

β GEOMETRIA PLANA (AFIN Y EUCLIDEA)

Es casi seguro que muchas, por no decir todas, las cuestiones que vamos a considerar aquí son ya conocidas de estudios anteriores, pues están en los programas de las enseñanzas medias. Esto nos permitirá avanzar con rapidez, pero no nos exime de abordar los problemas afines y euclídeos típicos de la geometría bidimensional, como son las intersecciones de rectas, el paralelismo, la distancia de un punto a una recta o el área de un triángulo.

11.3. PROBLEMAS AFINES (EN EL PLANO E_2)

Hablaremos aquí acerca de las rectas, sus ecuaciones, intersecciones de rectas y paralelismo. Según se ve, aquí las protagonistas son las rectas, que según dijimos recientemente en [202], son los conjuntos de puntos del tipo

$$P + \mathcal{V}(\bar{u}) = \{X = P + \lambda\bar{u} / \lambda \in \mathbb{R}\},$$

donde $P \in E_2$ es un punto fijo (por el que pasa la recta) y $\bar{u} \in \mathbb{R}^2$ es un vector dado (que define la dirección de la recta); la recta que pasa por dos puntos $P, Q \in E_2$ es el conjunto $P + \mathcal{V}(\overrightarrow{PQ})$, de los puntos $X = P + \lambda\overrightarrow{PQ}$ (para $\lambda \in \mathbb{R}$). Para representar a estas rectas, se suele poner:

$$X = P + \lambda\bar{u} \quad \text{o} \quad X = P + \lambda\overrightarrow{PQ} \quad (\text{donde } \lambda \in \mathbb{R})$$

ECUACIONES DE LAS RECTAS (DE E_2)

En el plano E_2 se adoptará una referencia cartesiana $(O; \bar{e}_1, \bar{e}_2)$, lo que permite asignar coordenadas (x_1, x_2) a todo punto $X \in E_2$. Vamos a hablar aquí sobre la relación a la que deben verificar x_1 y x_2 para que X pertenezca a una recta dada y recíprocamente. Una tal relación recibe el nombre de ecuación de la recta.

[204]

En el plano E_2 se adopta una referencia cartesiana cualquiera; respecto de ella:

1. La recta, del plano E_2, cuya ecuación vectorial es $X = P + \lambda\bar{u}$, o sea, la que pasa por el punto $P(p_1, p_2)$ y tiene la dirección del vector no nulo $\bar{u}(\bar{u}_1, \bar{u}_2)$, admite las siguientes ecuaciones con coordenadas:

$$\begin{cases} x_1 = p_1 + \lambda\bar{u} \\ x_2 = p_2 + \lambda\bar{u}_2 \end{cases} \quad (\text{para } \lambda \in \mathbb{R}) \quad \text{ó} \quad \frac{x_1 - p_1}{u_1} = \frac{x_2 - p_2}{u_2} \qquad [1]$$

(ecuaciones paramétricas) (ecuación continua)

2. La recta de E_2 que pasa por los puntos $P(p_1, p_2)$ y $Q(q_1, q_2)$ admite por ecuaciones a:

$$\begin{cases} x_1 = p_1 + \lambda(q_1 - p_1) \\ x_2 = p_2 + \lambda(q_2 - p_2) \end{cases} \quad (\text{para } \lambda \in \mathbb{R}) \quad \text{ó}$$

$$\frac{x_1 - p_1}{q_1 - p_1} = \frac{x_2 - p_2}{q_2 - p_2} \quad \text{ó} \quad \begin{vmatrix} x_1 & x_2 & 1 \\ p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \end{vmatrix} = 0 \qquad [2]$$

3. Toda recta de E_2 admite ecuación de la forma [3], que abajo se da, para ciertos números $a_0, a_1, a_2 \in \mathbb{R}$ con $(a_1, a_2) \neq (0, 0)$; recíprocamente, toda ecuación de dicho tipo representa a una recta de E_2. Si $P(p_1, p_2)$ es un punto de una recta, la ecuación general de ésta puede ponerse en la forma [4]:

$$\left.\begin{aligned} a_0 + a_1x_1 + a_2x_2 = 0 \quad \text{ó} \\ a_1(x_1 - p_1) + a_2(x_2 - p_2) = 0 \end{aligned}\right\} \begin{pmatrix} \text{ecuación general} \\ \text{de una recta de } E_2 \end{pmatrix} \quad \begin{matrix} [3] \\ [4] \end{matrix}$$

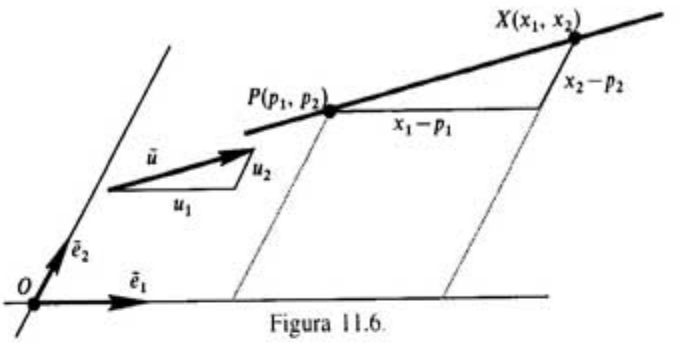


Figura 11.6.

COMPROBACIONES

1. Los puntos de la recta son aquellos $X(x_i)$ tales que $\overrightarrow{PX} = \lambda\bar{u}$ para algún $\lambda \in \mathbb{R}$, o sea, tales que $x_i - p_i = \lambda\bar{u}_i$ ($i = 1, 2$ y $\lambda \in \mathbb{R}$), que son las ecuaciones paramétricas. La ecuación continua equivale, obviamente, a las paramétricas; si algún u_i fuese nulo, véase lo que se dice más adelante, en las observaciones.
2. Aquí se repite lo anterior, tomando $\bar{u} = \overrightarrow{PQ}$. La última ecuación, la ecuación «en forma de determinante» es equivalente a la ecuación continua, lo que se ve fácilmente restando la 2.ª fila del determinante a cada una de las otras dos.
3. Según se acaba de comprobar, toda recta admite ecuaciones paramétricas [1], que equivalen a $u_2(x_1 - p_1) - u_1(x_2 - p_2) = 0$, que es de los tipos [3] ó [4]; nótese que los coeficientes de x_1 y x_2 no pueden ser los dos nulos, pues u_1 y u_2 son los componentes de un vector no nulo. Recíprocamente, dada una ecuación del tipo [3] con $(a_1, a_2) \neq (0, 0)$, por ejemplo, con $a_1 \neq 0$, es evidente que dicha ecuación lo es de la recta que pase por el punto $P(-a_0/a_1, 0)$ y tiene la dirección del vector no nulo $\bar{u}(a_2, -a_1)$.

OBSERVACIONES

- Si en las ecuaciones paramétricas [1] fuese $u_1 = 0$ o $u_2 = 0$, no tendría sentido la ecuación continua. No obstante, hay quien abusa de la notación poniendo un cero como denominador; en tal caso debe entenderse que se quiere indicar (de mala forma) que el correspondiente numerador es nulo.
- Normalmente, y sobre todo en las aplicaciones, a las coordenadas x_1 y x_2 de un punto genérico del plano E_2 se las suele llamar x e y. Así, por ejemplo, la ecuación general de una recta se suele expresar poniendo $ax + by + c = 0$, donde $(a, b) \neq (0, 0)$.
- Si en la ecuación general $ax + by + c = 0$. de una recta de E_2, es $b \neq 0$, la ecuación se puede poner en la forma $y = mx + h$, donde los números m y h se llaman pendiente y ordenada en el origen, de la recta dada.
- Si una recta corta a los ejes coordenados en los puntos de coordenadas $(\alpha, 0)$ y $(0, \beta)$, entonces la recta admite por ecuación a la $(x/\alpha) + (y/\beta) = 1$.

EJERCICIO

Hallar las ecuaciones generales de las rectas que tienen la dirección del vector $\bar{u}(a, b)$ y que determinan en los ejes coordenadas, OX y OY, segmentos OA y OB cuyas amplitudes suman la unidad.

RESOLUCION

Como la recta tiene la dirección de $\bar{u}$, su ecuación general es $bx - ay = h$, para cierta $h \in \mathbb{R}$. Como $A(h/b, 0)$ y $B(0, -h/a)$, ha de verificarse que:

$$\left|\frac{h}{b}\right| + \left|\frac{h}{a}\right| = 1, \quad \text{luego} \quad |h| = \frac{|ab|}{|a| + |b|}$$

Por tanto, hay dos rectas solución, que son las de ecuaciones:

$$bx - ay = \pm \frac{|ab|}{|a| + |b|}$$

PARALELISMO

[205]

Dos rectas se dicen paralelas si tienen la misma dirección; esto es, las rectas $P + V(\bar{u})$ y $Q + V(\bar{v})$ son paralelas si $\bar{u}$ y $\bar{v}$ son proporcionales. Si en el plano E_2 se adopta una referencia cartesiana cualquiera, respecto de ella:

1. La relación $u_1/v_1 = u_2/v_2$ (*) es necesaria y suficiente para el paralelismo de:

$$(x_1, x_2) = (p_1, p_2) + \lambda(u_1, u_2) \quad \text{y} \quad (x_1, x_2) = (q_1, q_2) + \lambda(v_1, v_2)$$

2. La relación $a_1/b_1 = a_2/b_2$ (*) es necesaria y suficiente para el paralelismo de:

$$a_0 + a_1x_1 + a_2x_2 = 0 \quad \text{y} \quad b_0 + b_1x_1 + b_2x_2 = 0$$

3. La relación $a_1\bar{u}_1 + a_2\bar{u}_2 = 0$ es necesaria y suficiente para el paralelismo de:

$$a_0 + a_1x_1 + a_2x_2 = 0 \quad \text{y} \quad (x_1, x_2) = (p_1, p_2) + \lambda(u_1, u_2)$$

(*) Si un denominador es nulo, esta relación significará que también ha de ser nulo el correspondiente numerador.

COMPROBACIONES

1. Las dos rectas dadas son paralelas si y sólo si sus vectores de dirección $\bar{u}(u_1, u_2)$ y $\bar{v}(v_1, v_2)$ son proporcionales, esto es, si $u_1/v_1 = u_2/v_2$.
2. Las dos rectas dadas tienen por vectores de dirección a los $\bar{u}(a_2, -a_1)$ y $\bar{v}(b_2, -b_1)$ y, por ello, el paralelismo equivale a $a_2/b_2 = a_1/b_1$.
3. Las dos rectas dadas tienen por vectores de dirección a los $\bar{u}(u_1, u_2)$ y $\bar{v}(a_2, -a_2)$, luego el paralelismo equivale a $u_1/a_2 = u_2/-a_1$, o sea, a $a_1u_1 + a_2u_2 = 0$

EJERCICIO

Dado el punto $P(p_1, p_2)$ y las rectas r y s de ecuaciones:

$$r = \begin{cases} x_1 = p_1 + \lambda u_1 \\ x_2 = p_2 + \lambda u_2 \end{cases} \quad (\lambda \in \mathbb{R})$$

y

$$s = a_0 + a_1x_1 + a_2x_2 = 0$$

hallar ecuaciones de las rectas, r' y s', paralelas a r y a s por P.

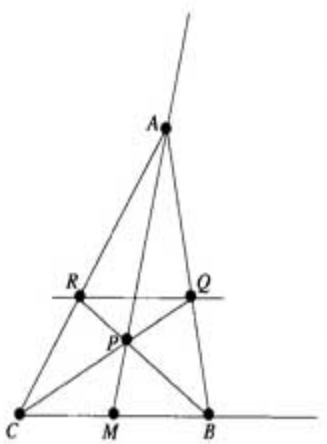
RESOLUCION

De acuerdo con la definición de paralelismo, es obvio que

$$r' = \begin{cases} x_1 = p_1 + \lambda u_1 \\ x_2 = p_2 + \lambda u_2 \end{cases} \quad (\lambda \in \mathbb{R})$$

y

$$s' = a_1(x_1 - p_1) + a_2(x_2 - p_2) = 0$$



EJERCICIO

Se dan tres puntos no alineados A, B y C; sea M el punto medio del segmento BC. En la recta AM se considera un punto cualquiera P; sea Q la intersección de AB y CP; sea R la intersección de AC y BP. Pruébese que QR es paralela a BC.

RESOLUCION

Tomemos la referencia cuyo origen es M y cuyos vectores son $\overrightarrow{MB}$ y $\overrightarrow{MA}$ (llamemos x e y a las coordenadas). En esta referencia $A(0, 1)$, $B(1, 0)$, $C(-1, 0)$

y $P(0, a)$, para $a \in \mathbb{R}$ cualquiera. Las ecuaciones de las rectas del enunciado son:

$$AB: x+y=1 \quad BP: ax+y=a$$
$$AC: -x+y=1 \quad CP: -ax+y=a$$

Por tanto, los puntos Q y R y la recta QR son:

$$Q\left(\frac{1-a}{1+a}, \frac{2a}{1+a}\right) \quad R\left(\frac{a-1}{1+a}, \frac{2a}{1+a}\right) \quad QR: y=\frac{2a}{1+a}$$

Ha resultado, pues, que QR es paralela al eje $y=0$, que es BC.

INTERSECCIONES DE RECTAS (DE E_2)

[206]

En el plano E_2 y respecto de cualquier referencia cartesiana que se tome en él, se verifica:

INTERSECCION DE DOS RECTAS: Dadas dos rectas r y s cuyas ecuaciones generales son $r: a_1x_1 + a_2x_2 + a_0 = 0$ y $s: b_1x_1 + b_2x_2 + b_0 = 0$ y llamando M y M' a las matrices:

$$M = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \quad \text{y} \quad M' = \begin{bmatrix} a_1 & a_2 & a_0 \\ b_1 & b_2 & b_0 \end{bmatrix}$$

- [rang $M = 2$] $\Rightarrow$ $r \cap s$ consta de un solo punto (r y s son secantes).
- [rang $M = 1$ y rang $M' = 2$] $\Rightarrow$ $r \cap s = \varnothing$ y las rectas son paralelas.
- [rang M = rang $M' = 1$] $\Rightarrow$ las rectas coinciden (son la misma).

INTERSECCION DE TRES RECTAS: Dadas las siguientes rectas r, s y t y llamando M y M' a las matrices que abajo se señalan

$$r: a_1x_1 + a_2x_2 + a_0 = 0$$
$$s: b_1x_1 + b_2x_2 + b_0 = 0$$
$$t: c_1x_1 + c_2x_2 + c_0 = 0$$

$$M = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \quad ; \quad M' = \begin{bmatrix} a_1 & a_2 & a_0 \\ b_1 & b_2 & b_0 \\ c_1 & c_2 & c_0 \end{bmatrix}$$

- [rang $M' = 3$] $\Rightarrow$ $r \cap s \cap t = \varnothing$ y forman triángulo o dos son paralelas.
- [rang M = rang $M' = 2$] $\Rightarrow$ $r \cap s \cap t$ consta de un solo punto.
- [rang $M = 1$ y rang $M' = 2$] $\Rightarrow$ $r \cap s \cap t = \varnothing$ (son paralelas y no coinciden las tres).
- [rang M = rang $M' = 1$] $\Rightarrow$ las tres rectas coinciden (son la misma).

COMPROBACIONES

Un punto $X(x_1, x_2)$ pertenecerá a la intersección de las rectas dadas si y sólo si, (x_1, x_2) es solución del sistema que forman las ecuaciones de dichas rectas. Al recurrir al teorema de Rouché, relativo a la discusión de un sistema de ecuaciones lineales, se obtienen directamente los resultados del enunciado.

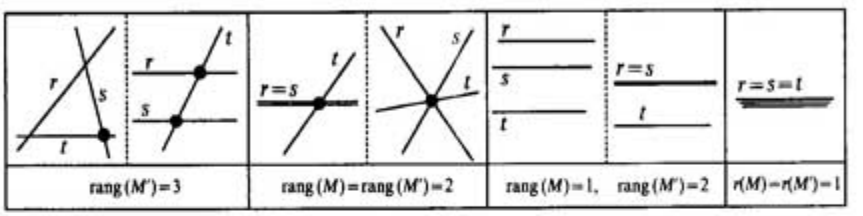


Figura 11.7.

EJERCICIO

Estudiar la posición relativa de las tres rectas siguientes:

$$r: x + y = 1 \quad ; \quad s: 2x + 3y = 1 \quad ; \quad t: 3x + \alpha y = 2$$

en función del valor del parámetro $\alpha \in \mathbb{R}$.

RESOLUCION

Recurriendo a las matrices M y M' del sistema de las ecuaciones de las tres rectas:

$$M = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & \alpha \end{bmatrix} \quad ; \quad M' = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 3 & \alpha & 2 \end{bmatrix}$$

$$\text{rang } M = 2$$

$$\text{rang } M' = \begin{cases} 3, & \text{si} \quad \alpha \neq 4 \\ 2, & \text{si} \quad \alpha = 4, \end{cases}$$

se concluye fácilmente que:

- Para $\alpha = 4$, es rang M = rang $M' = 2$; en este caso r, s y t se cortan las tres (y dos a dos) en un punto, que es el $P(2, -1)$.
- Para $\alpha \neq 4$, es rang $M' = 3$; r, s y t no tienen ningún punto en común. Para $\alpha = 3$, r y t son paralelas y distintas y s es secante con ellas. Para $\alpha = 9/2$, s y t son paralelas y distintas y r es secante con ellas. Para $\alpha \neq 4, 3, 9/2$, las tres rectas forman triángulo.

HACES DE RECTAS (DE E_2)

[207]

Respecto de una referencia cartesiana cualquiera, en el plano E_2, se tiene:

Dadas dos rectas $r: a_1x_1 + a_2x_2 + a_0 = 0$ y $s: b_1x_1 + b_2x_2 + b_0 = 0$, distintas, se llama *haz de rectas*, que determinan r y s, al conjunto de todas las rectas que tienen ecuación del siguiente tipo:

$$\alpha(a_1x_1 + a_2x_2 + a_0) + \beta(b_1x_1 + b_2x_2 + b_0) = 0$$

para cualesquiera $\alpha, \beta \in \mathbb{R}$ tales que $(\alpha,\beta) \neq (0, 0)$ ó

$$(a_1x_1 + a_2x_2 + a_0) + \lambda(b_1x_1 + b_2x_2 + b_0) = 0$$

para $\lambda \in \overline{\mathbb{R}}$ (*). Este haz está formado por las siguientes rectas:

1. Si r y s se cortan, en un punto P_0, entonces el haz lo forman todas las rectas que pasen por P_0.
2. Si r y s son paralelas, el haz lo forman todas las rectas paralelas a las r y s.

(*) Se toma $\lambda = \beta/\alpha$. Esta forma de dar el haz es defectuosa, pues no incluye a la recta s, que se obtiene para $(\alpha, \beta) = (0, 1)$. Para intentar remediarlo se admite que s es el límite de la recta del haz para $\lambda \to \infty$.

COMPROBACION

Para simplificar la notación, pongamos $a(x)$ y $b(x)$ para representar a $a_1x_1 + a_2x_2 + a_0$ y $b_1x_1 + b_2x_2 + b_0$. Por cualquier punto P del plano, excepto el posible punto de intersección de r y s, pasa una y sólo una recta del haz, que es la que se obtiene para $\alpha = b(p)$ y $\beta = -a(p)$.

1. Si r y s se cortan en un punto P_0, esto es, si $a(p_0) = 0$ y $b(p_0) = 0$, entonces todas las rectas del haz pasan por P_0 ya que $\alpha a(p_0) + \beta b(p_0) = 0$ para cualesquiera α y β. Recíprocamente, si t es una recta que pasa por P_0, entonces t ha de ser una de las rectas del haz, ya que, tomando un punto P de ella, distinto de P_0 y según se dijo antes, por P pasa una sola recta t' del haz y, como t y t' pasan ambas por P y P_0, ha de ser $t' = t$, luego t es del haz.
2. Si r y s son paralelas, esto es, si (a_1, a_2) y (b_1, b_2) son proporcionales, entonces todas las rectas del haz son paralelas a ellas, pues el par $(\alpha a_1 + \beta b_1, \alpha a_2 + \beta b_2)$ es ahora proporcional a los (a_1, a_2) y (b_1, b_2). Recíprocamente, si t es una recta paralela a las r y s, entonces t ha de ser una de las rectas del haz, ya que, tomando un punto P cualquiera de ella y según se obtuvo anteriormente, por P pasa una sola recta t' del haz y, como t y t' son paralelas a las r y s y pasan por P, ha de ser $t' = t$, luego t es del haz.

EJERCICIO

Hallar la recta t que es paralela a r y pasa por la intersección de s_1 y s_2, siendo r, s_1 y s_2 las rectas de ecuaciones generales

$$r: 3x + 2y + 5 = 0 \quad ; \quad s_1: x + 2y + 1 = 0 \quad ; \quad s_2: 2x + y = 7$$

RESOLUCION

La recta t, como pasa por la intersección de s_1 y s_2, pertenece al haz que éstas definen, luego

$$t: (x + 2y + 1) + \lambda(2x + y - 7) = 0 \quad \text{para un cierto } \lambda \in \overline{\mathbb{R}}$$

Dado que t y r son paralelas, tendrán proporcionales los coeficientes de x e y:

$$\frac{1 + 2\lambda}{3} = \frac{2 + \lambda}{2}, \quad \text{luego} \quad \lambda = 4 \quad \text{y} \quad t: 3x + 2y = 9$$

11.4. PROBLEMAS EUCLIDEOS (EN EL PLANO E_2)

Suponemos ahora que el plano E_2 es euclídeo, esto es, dispondremos del producto escalar entre vectores libres (vectores de $\mathbb{R}^2$). Siempre que, aquí, precisemos de una referencia cartesiana, la tomaremos rectangular, con lo que simplificaremos notablemente las expresiones que proporcionan las distancias y los ángulos. Así, en particular, en coordenadas rectangular la distancia del punto $X(x_1, x_2)$ al punto $Y(y_1, y_2)$ valdrá $d(X, Y) = [(y_1 - x_1)^2 + (y_2 - x_2)^2]^{1/2}$.

EXPRESION EUCLIDEA DE UNA RECTA (DE E_2)

[208]

Una recta r cualquiera, del plano euclídeo E_2, queda definida si se conoce un punto P de ella y un vector no nulo $\bar{a} \in \mathbb{R}^2$ ortogonal a su dirección. En función de P y $\bar{a}$, la recta r es el lugar geométrico de los puntos $X \in E_2$ tales que:

$$\bar{a} \cdot \overrightarrow{PX} = 0 \quad \text{(ecuación euclídea de la recta)}$$

Recurriendo a una referencia cartesiana rectangular y si en ella $P(p_1, p_2)$ y $\bar{a}(a_1, a_2)$, la anterior recta, $\bar{a} \cdot \overrightarrow{PX} = 0$, admite por ecuación a la:

$$a_1(x_1 - p_1) + a_2(x_2 - p_2) = 0$$

Así, pues, el vector $\bar{a}(a_1, a_2)$ es ortogonal (*) a la recta de ecuación $a_1x_1 + a_2x_2 + a_0 = 0$.

(*) Abusando del lenguaje, un vector se dice ortogonal o perpendicular a una recta si lo es a su dirección.

COMPROBACION

Dados $P \in E_2$ y $\bar{a} \in \mathbb{R}^2$, *si* $\bar{u}$ es un vector no nulo ortogonal a $\bar{a}$ y como $(\bar{a}, \bar{u})$ es una base de $\mathbb{R}^2$, decir que $\bar{a} \cdot \overrightarrow{PX} = 0$, esto es, que $\overrightarrow{PX}$ es ortogonal a $\bar{a}$, equivale a decir que $\overrightarrow{PX}$ tiene la dirección de $\bar{u}$, o sea, que $\overrightarrow{PX} = \lambda\bar{u}$ para algún $\lambda \in \mathbb{R}$, lo que significa que X recorre una recta (la que pasa por P y tiene la dirección de $\bar{u}$), como había que verificar.

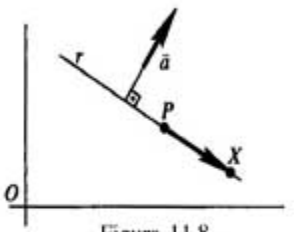


Figura 11.8.

Echando mano de la referencia rectangular, como en ella $\bar{a}(a_1, a_2)$ y $\overrightarrow{PX}(x_1 - p_1, x_2 - p_2)$, resulta que la relación $\bar{a} \cdot \overrightarrow{PX} = 0$ se expresa en la forma $a_1(x_1 - p_1) + a_2(x_2 - p_2) = 0$.

EJERCICIO

Hállese la ecuación de la altura correspondiente al vértice A del triángulo ABC, siendo (en coordenadas rectangulares):

$$A(2, 1); \quad B(1, 3); \quad C(5, 0)$$

RESOLUCION

El vector $\overrightarrow{BC}(4, -3)$ es perpendicular a la recta pedida, luego ésta es la $\overrightarrow{AX} \cdot \overrightarrow{BC} = 0$ (X = punto genérico de la recta), esto es:

$$(x_1 - 2)4 + (x_2 - 1)(-3) = 0, \quad \text{o sea} \quad 4x_1 - 3x_2 = 5$$

ANGULO DE DOS RECTAS

Conviene empezar notando que, dadas dos rectas r y s, del plano euclídeo E_2, con dos vectores $\bar{u}_r$ y $\bar{u}_s$, de las direcciones de r y s, se forman cuatro ángulos, uno de ellos, al que llamaremos α, está entre 0 y $\pi/2$; en función de él, los cuatro ángulos son:

$$\text{áng}\,(\bar{u}_r, \bar{u}_s) = \text{áng}\,(-\bar{u}_r, -\bar{u}_s) = \alpha \in [0, \pi/2]$$

$$\text{áng}\,(\bar{u}_s, \bar{u}_r) = \text{áng}\,(-\bar{u}_s, -\bar{u}_r) = -\alpha \in [-\pi/2, 0]$$

$$\text{áng}\,(\bar{u}_r, -\bar{u}_s) = \text{áng}\,(-\bar{u}_r, \bar{u}_s) = \alpha - \pi \in [-\pi, \pi/2]$$

$$\text{áng}\,(\bar{u}_s, -\bar{u}_r) = \text{áng}\,(-\bar{u}_s, \bar{u}_r) = \pi - \alpha \in [-\pi/2, \pi]$$

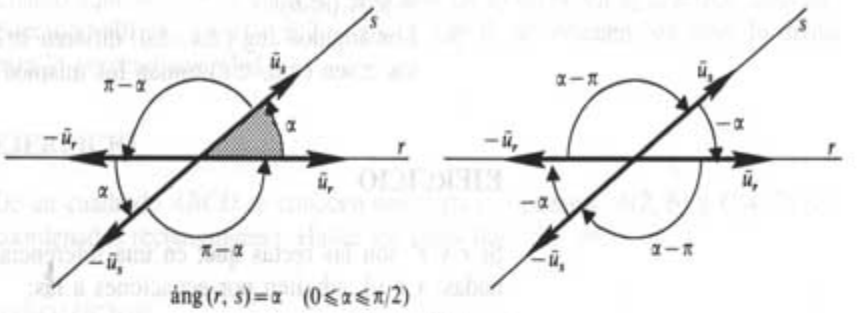


Figura 11.9.

Estos cuatro ángulos dependen sólo de las rectas r y s, esto es, ellos son los mismos si se sustituyen $\vec{u}_r$ y $\vec{u}_s$ por otros vectores, no nulos, de las direcciones de r y de s.

[209]

Se llama *ángulo* áng (r, s), de dos rectas r y s del plano euclídeo E_2, al ángulo comprendido entre 0 y $\pi/2$ de entre los cuatro ángulos que se pueden formar con un vector (no nulo) de la dirección de una de las rectas y otro de la dirección de la otra.
El áng (r, s) queda determinado por su coseno, que vale (*):

1. Si $\vec{u}=$ y $\vec{v}$ son vectores de las direcciones de r y s, entonces $\cos(r, s) = |\cos(\vec{u}, \vec{v})|$.
2. Si $\vec{a}$ y $\vec{b}$ son vectores ortogonales a las rectas r y s, entonces $\cos(r, s) = |\cos(\vec{a}, \vec{b})|$.
3. Si $\vec{u}$ tiene la dirección de r y $\vec{a}$ es ortogonal a s, entonces $\operatorname{sen}(r, s) = |\cos(\vec{u}, \vec{a})|$.

(*) Recuérdese que el coseno del ángulo de dos vectores $\vec{x}$ e $\vec{y}$ (no nulos) vale $\cos(\vec{x}, \vec{y}) = (\vec{x} \cdot \vec{y}) : (\|\vec{x}\| \, \|\vec{y}\|)$.

COMPROBACION

Según se dijo en la introducción anterior, dos rectas cualesquiera forman un único ángulo, comprendido entre 0 y $\pi/2$. Las expresiones que dan el coseno, o el seno, del ángulo áng (r, s) son correctas ya que, respectivamente:

1. Los cos $(\pm\vec{u}, \pm\vec{v})$ y cos $(\pm\vec{v}, \pm\vec{u})$ todos tienen el mismo valor absoluto y, de entre ellos, los que son positivos son iguales al cos (r, s).

2. Los cuatro ángulos que se forman con los vectores $\pm\bar{a}$ y $\pm\bar{b}$ son los mismo que los que se forman con $\pm\bar{u}$ y $\pm\bar{v}$, con lo que $|\cos(\bar{u}, \bar{v})| = |\cos(\bar{a}, \bar{b})|$.
3. Los ángulos áng $(\pm\bar{u}, \pm\bar{a})$ difieren $\pi/2$ de los áng $(\pm\bar{u}, \pm\bar{v})$; por ello, los $\pm$sen $(\pm\bar{u}, \pm\bar{v})$ toman los mismos valores que los $\pm\cos(\bar{u}, \bar{a})$.

EJERCICIO

Si r y r' son las rectas que, en una referencia cartesiana rectangular (coordenadas: x e y), admiten por ecuaciones a las:

$$r: y = mx + h \quad ; \quad r': y = m'x + h'$$

compruébese que

$$\text{tg}\,(r, r') = \left|\frac{m - m'}{1 + mm'}\right|$$

RESOLUCION

Los vectores $\bar{u}(1, m)$ y $\bar{u}'(1, m')$ tienen las direcciones de r y r' y, por ello:

$$\cos(r, r') = |\cos(\bar{u}, \bar{u}')| = \frac{|1 + mm'|}{\sqrt{1 + m^2}\sqrt{1 + m'^2}}$$

$$\text{tg}\,(r, r') = \frac{\sqrt{1 - \cos^2(r, r')}}{\cos(r, r')} = \frac{|m - m'|}{|1 + mm'|}$$

RECTAS PERPENDICULARES

[210]

Dos rectas r y s, del plano euclídeo E_2, se dicen *perpendiculares* si el ángulo que forman es recto, lo que equivale a que sea $\bar{u} \cdot \bar{v} = 0$ o a que sea $\bar{a} \cdot \bar{b} = 0$, donde $\bar{u}$ y $\bar{v}$ son vectores (no nulos) de las direcciones de r y s y donde $\bar{a}$ y $\bar{b}$ son vectores (no nulos) ortogonales a r y a s.

Si se adopta una referencia rectangular en E_2, entonces: 1) las rectas perpendiculares a una que tienen la dirección del vector (u_1, u_2), tienen la dirección del vector $(u_2, -u_1)$; y 2) las rectas perpendiculares a la de ecuación $a_1x_1 + a_2x_2 + h = 0$ tienen ecuación del tipo $a_2x_1 - a_1x_2 + k = 0$.

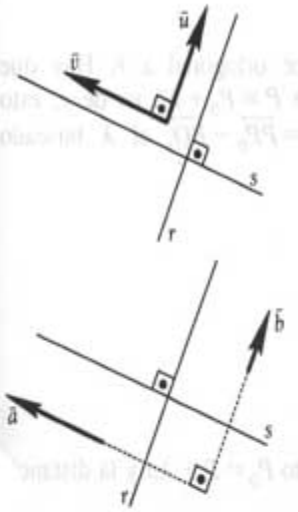


Figura 11.10.

COMPROBACION

Cuanto aquí se dice es un caso particular de lo dicho en el apartado anterior. Tomando allí áng $(r, s) = \pi/2$ y cos $(r, s) = 0$, se obtienen, sin más, lo dicho para la perpendicularidad.

EJERCICIO

De un cuadrado $ABCD$ se conocen dos vértices opuestos, $A(2, 6)$ y $C(4, 2)$ (en coordenadas rectangulares). Hallar los otros dos vértices.

RESOLUCION

Los vectores $\overrightarrow{AC}$ y $\overrightarrow{BD}$ tienen igual longitud y son perpendiculares; como $\overrightarrow{AC}(2, -4)$, será $\overrightarrow{BD}(4, 2)$ o $\overrightarrow{BD}(-4, 2)$. El punto medio de AB, que es $O(3, 4)$, es también el punto medio de BD y, por tanto, B y D son los puntos $O \pm \frac{1}{2}\overrightarrow{BD}$; así, pues $B(5, 5)$ y $D(1, -3)$.

DISTANCIA DE UN PUNTO A UNA RECTA

[211]

Sean dados, en el plano euclídeo E_2, un punto P y una recta r. Si r no pasa por P, entonces existe un punto P_0 de r tal que el vector $\overrightarrow{PP_0}$ es ortogonal a r; se dice que P_0 es la *proyección ortogonal* de P sobre r. La distancia de P a P_0 es la menor de las distancias de P a los puntos de r y se le llama *distancia* del punto P a la recta r.

Si, respecto de una referencia rectangular, $P(p_1, p_2)$ y r tiene por ecuación a $a_1x_1 + a_2x_2 + a_0 = 0$, entonces la distancia de P a r vale:

$$d(P, r) = \frac{|a_1p_1 + a_2p_2 + a_0|}{\sqrt{a_1^2 + a_2^2}}$$

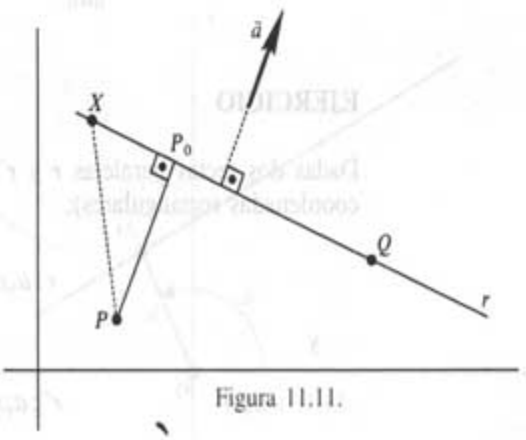


Figura 11.11.

DEMOSTRACION

Llamemos Q a un punto de r y $\bar{a}$ a un vector ortogonal a r. Hay que comprobar que existe un $\lambda \in \mathbb{R}$ tal que el punto $P = P_0 + \lambda\bar{a}$ es de r, esto es, se verifica que $\overrightarrow{QP_0} \cdot \bar{a} = 0$. Poniendo $\overrightarrow{QP_0} = \overrightarrow{PP_0} - \overrightarrow{PQ}$, el λ buscado es, pues, aquel que verifica a

$$(\lambda\bar{a} - \overrightarrow{PQ}) \cdot \bar{a} = 0,$$

luego

$$\lambda = \frac{\overrightarrow{PQ} \cdot \bar{a}}{\|\bar{a}\|^2}$$

La proyección ortogonal de P sobre r es el punto $P_0 = P + \lambda\bar{a}$ y la distancia $d(P, P_0)$ será:

$$d(P, P_0) = \|\lambda\bar{a}\| = |\lambda|\,\|\bar{a}\| = \frac{|\overrightarrow{PQ} \cdot \bar{a}|}{\|\bar{a}\|} \qquad [1]$$

Para cualquiera que sea el punto $X \neq P_0$ de r, como el triángulo PP_0X es rectángulo en P_0 y según el teorema de Pitágoras (véase [200]), se verifica que

$$\|\overrightarrow{PX}\|^2 = \|\overrightarrow{PP_0}\|^2 + \|\overrightarrow{P_0X}\|^2 > \|\overrightarrow{PP_0}\|^2,$$

luego

$$\|\overrightarrow{PX}\| > \|\overrightarrow{PP_0}\|$$

Por tanto, $d(P, P_0) = \|\overrightarrow{PP_0}\|$ es la menor de las distancias de P a los puntos de r.

Finalmente, al recurrir a coordenadas rectangulares, como puede tomarse $\bar{a}(a_1, a_2)$ y es $P(p_1, p_2)$ y teniendo en cuenta que, por ser $Q(q_1, q_2)$ un punto de r, se verifica que $a_1q_1 + a_2q_2 = -a_0$, se obtiene que:

$$d(P, r) = d(P, P_0) = \frac{|\overrightarrow{PQ} \cdot \bar{a}|}{\|\bar{a}\|} = \frac{|(q_1 - p_1)a_1 + (q_2 - p_2)a_2|}{\sqrt{a_1^2 + a_2^2}} = \frac{|a_1p_1 + a_2p_2 + a_0|}{\sqrt{a_1^2 + a_2^2}}$$

EJERCICIO

Dadas dos rectas paralelas r y r', cuyas ecuaciones serán, pues, del tipo (en coordenadas rectangulares):

$$r: a_1x_1 + a_2x_2 + a_0 = 0$$

y

$$r': a_1x_1 + a_2x_2 + a_0' = 0$$

pruébese que la distancia de un punto cualquiera de r a r' es constante. A esta distancia se la llama distancia de r a r' y vale:

$$d(r, r') = \frac{|a_0 - a_0'|}{\sqrt{a_1^2 + a_2^2}}$$

RESOLUCION

Sea $P(p_1, p_2)$ un punto cualquiera de r, esto es, tal que $a_1p_1 + a_2p_2 + a_0 = 0$; la distancia de P a r' vale, según se acaba de probar:

$$d(P, r') = \frac{|a_1p_1 + a_2p_2 + a_0'|}{\sqrt{a_1^2 + a_2^2}} = \frac{|-a_0 + a_0'|}{\sqrt{a_1^2 + a_2^2}}$$

que es constante (no depende de quien sea $P \in r$) y coincide con el valor que se da en el enunciado.

ECUACION NORMAL DE UNA RECTA

[212]

Sea r una recta del plano euclídeo E_2, en el que se ha adoptado una referencia cartesiana rectangular $(O; \bar{e}_1, \bar{e}_2)$. Llamando O_r a la proyección ortogonal de O sobre r y si $\overrightarrow{OO_r}$ $(d \cos \theta, d \operatorname{sen} \theta)$, donde $d > 0$ es la distancia de O a r, entonces r admite por ecuación a:

$$x_1 \cos \theta + x_2 \operatorname{sen} \theta = d$$

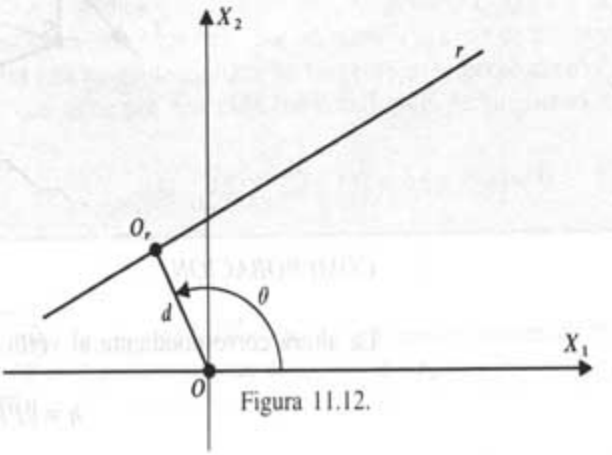


Figura 11.12.

COMPROBACION

Un punto X es de r si y sólo si $\overrightarrow{O_rX} \cdot \overrightarrow{OO_r} = 0$, esto es, si:

$$(x_1 - d\cos\theta)d\cos\theta + (x_2 - d\,\text{sen}\,\theta)d\,\text{sen}\,\theta = 0$$

que es la ecuación que había que obtener.

AREA DE UN TRIANGULO

[213]

Supongamos que, en el plano euclídeo E_2, se dan tres puntos no alineados $P(p_1, p_2)$, $Q(q_1, q_2)$ y $R(r_1, r_2)$ (las coordenadas son rectangulares). El área del triángulo PQR vale:

$$\text{Area } \hat{PQR} = \tfrac{1}{2}\|\overrightarrow{PQ} \wedge \overrightarrow{PR}\| = \tfrac{1}{2}\left|\det \begin{bmatrix} p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \\ r_1 & r_2 & 1 \end{bmatrix}\right| \quad (*)$$

(*) Las barras verticales son de «valor absoluto». Para el producto vectorial $\overrightarrow{PQ} \wedge \overrightarrow{PR}$, los vectores $\overrightarrow{PQ}$ y $\overrightarrow{PR}$, que son de $\mathbb{R}^2$, se supone que son de $\mathbb{R}^3$ con su tercera componente nula.

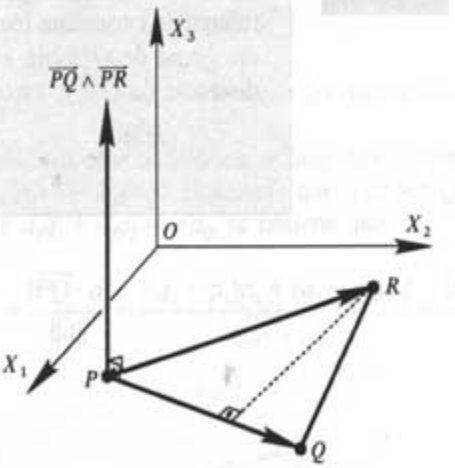


Figura 11.13.

COMPROBACION

La altura correspondiente al vértice R se puede expresar en la forma:

$$h = \|\overrightarrow{PR}\|\,\text{sen}\,(\overrightarrow{PQ}, \overrightarrow{PR})$$

Recurriendo a la norma de un producto vectorial (véase [158], II), el área del triángulo, que vale $(1/2)\,\|\overrightarrow{PQ}\|h$, se puede poner:

$$\text{Area } \overset{\triangle}{PQR} = \tfrac{1}{2}\|\overrightarrow{PQ}\|h = \tfrac{1}{2}\|\overrightarrow{PQ}\|\,\|\overrightarrow{PR}\|\,\text{sen}\,(\overrightarrow{PQ}, \overrightarrow{PR}) = \tfrac{1}{2}\|\overrightarrow{PQ}\wedge\overrightarrow{PR}\| \ (*) \qquad [1]$$

El vector $\overrightarrow{PQ}\wedge\overrightarrow{PR}$ tiene nulos sus dos primeros componentes (véase [160]) y su tercera componente vale

$$\det\begin{bmatrix} q_1-p_1 & q_2-p_2 \\ r_1-p_1 & r_2-p_2 \end{bmatrix} = \det\begin{bmatrix} p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \\ r_1 & r_2 & 1 \end{bmatrix}$$

(la última igualdad se comprueba restando, en el último determinante, la primera fila a cada una de las otras dos y, después, desarrollando por la tercera columna). Llevando estos valores a la expresión [1], se obtiene la fórmula del enunciado.

LAS CIRCUNFERENCIAS

[214]

(I) En el plano euclídeo E_2, se llama *circunferencia*, de centro en el punto $C \in E_2$ y cuyo radio es $\rho > 0$, al lugar geométrico de los puntos de E_2 que distan ρ de C. Adoptando una referencia cartesiana rectangular en E_2 y si $C(a, b)$, dicha circunferencia está formada por los puntos $X(x, y)$ tales que:

$$(x-a)^2+(y-b)=\rho^2 \quad \text{(ec. explícita)} \qquad \text{o} \qquad \begin{cases} x = a+\rho\,\cos\,\theta \\ y = b+\rho\,\text{sen}\,\theta \end{cases} \quad (\text{para } 0 \leq \theta < 2\pi)$$

(II) La circunferencia anterior y una recta r se cortan en dos, un o ningún punto, según que la distancia d de C a r sea $d<\rho$, $d=\rho$ o $d>\rho$ y se dice que r es secante, *tangente* o exterior a la circunferencia, respectivamente. Por un punto $P(x_0, y_0)$ de la circunferencia pasa una tangente a ella; esta tangente es la perpendicular a CP por P, que tiene por ecuación (en coordenada rectangulares x, y) a:

$$(x_0-a)(x-x_0)+(y_0-b)(y-y_0)=0$$

(*) Para $\overrightarrow{PQ}(q_1-p_1, q_2-p_2)$ y $\overrightarrow{PR}(r_1-p_1, r_2-p_2)$, que son vectores de $\mathbb{R}^2$, convendremos en que $\overrightarrow{PQ}\wedge\overrightarrow{PR}$ es el producto vectorial de $(q_1-p_1, q_2-p_2, 0)$ y $(r_1-p_1, r_2-p_2, 0)$, que son vectores de $\mathbb{R}^3$.

DEMOSTRACION

(I) Un punto $X(x, y)$ será de la circunferencia si y sólo si $\|\overrightarrow{CX}\| = \rho$, esto es, si $\|\overrightarrow{CX}\|^2 = \rho^2$; como $\overrightarrow{CX}(x-a, y-b)$, la última relación se puede poner en la forma:

$$(x-a)^2 + (y-b)^2 = \rho^2$$

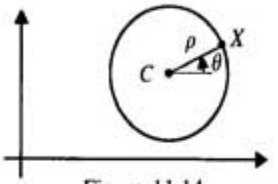


Figura 11.14.

También se puede decir que X pertenece a la circunferencia si y sólo si es $X = C + \rho\bar{u}$, donde $\bar{u} \in \mathbb{R}^2$ es un vector unitario cualquiera, esto es $\bar{u} = (\cos\theta, \text{sen}\ \theta)$ para $\theta \in [0, 2\pi]$. Por ello, la circunferencia admite las siguientes ecuaciones paramétricas

$$\begin{cases} x = a + \rho\cos\theta \\ y = b + \rho\ \text{sen}\ \theta \end{cases} \quad \text{para} \quad 0 \leq \theta < 2\pi$$

(II) Llamando P a la proyección ortogonal de C sobre r y $\bar{u}$ si es un vector unitario de la dirección de r, los puntos de esta recta r serán los $X = P + \lambda\bar{u}$ para $\lambda \in \mathbb{R}$. Un punto $X \in r$ será también de la circunferencia si y sólo si $\|\overrightarrow{CX}\| = \rho$; como el triángulo CPX es rectángulo en P y según el teorema de Pitágoras (véase [200]), la última relación es equivalente a

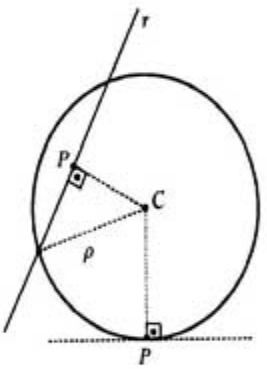


Figura 11.15.

$$\rho^2 = \|\overrightarrow{CX}\|^2 = \|\overrightarrow{CP}\|^2 + \|\overrightarrow{PX}\|^2 = d^2 + \lambda^2 \quad \text{o} \quad \lambda^2 = \rho^2 - d^2$$

Por tanto, según que sea $\rho > d$, $\rho = d$ o $\rho < d$, habrá dos, una o ninguna solución para λ, esto es, dos, un o ningún punto $(X = P + \lambda\bar{u})$ de intersección. Nótese que el caso de tangencia (la intersección es un punto único) se presenta cuando y sólo cuando la distancia de C a r es ρ y, en tal caso, el punto P de tangencia (intersección) es la proyección ortogonal de C sobre r, de manera que $\overrightarrow{CP}$ es ortogonal a r.

Dado el punto $P(x_0, y_0)$ de la circunferencia, la tangente en él es pues, la recta que pasa por P y es ortogonal a $\overrightarrow{CP}(x_0 - a, y_0 - b)$, que tiene por ecuación a

$$(x_0 - a)(x - x_0) + (y_0 - b)(y - y_0) = 0$$

EJERCICIO

En el plano euclídeo E_2 y en coordenadas rectangulares (x, y), se considera la circunferencia de centro en $C(a, b)$ y cuyo radio es ρ. Si $P(x_0, y_0)$ es un punto exterior a la circunferencia, obtener la ecuación conjunta de las dos tangentes que se pueden trazar desde P a la circunferencia.

RESOLUCION

Como las tangentes a la circunferencia son las rectas que están a distancia ρ de C, resulta que un punto $X(\alpha, \beta)$ será tal que la recta PX es tangente a la

circunferencia si y sólo si $d(C, PX) = \rho$. Puesto que la recta PX tiene por ecuación a

$$\frac{x - x_0}{\alpha - x_0} = \frac{y - y_0}{\beta - y_0} \quad \text{o} \quad (\beta - y_0)(x - x_0) - (\alpha - x_0)(y - y_0) = 0$$

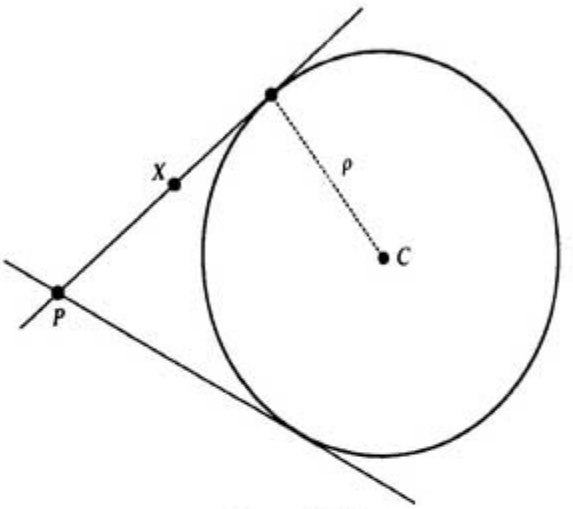


Figura 11.16.

la condición $d(C, PX) = \rho$ tomará la forma:

$$\rho = \frac{|(\beta - y_0)(a - x_0) - (\alpha - x_0)(b - y_0)|}{\sqrt{(\beta - y_0)^2 + (\alpha - x_0)^2}}$$

Elevando al cuadrado y operando, la anterior relación queda:

$$(\alpha - x_0)^2[\rho^2 - (b - y_0)^2] + (\beta - y_0)^2[\rho^2 - (a - x_0)^2] +$$
$$+ 2(\alpha - x_0)(\beta - y_0)(a - x_0)(b - y_0) = 0$$

Por tanto, esta última ecuación es la que deben satisfacer los puntos $X(\alpha, \beta)$ que recorren las tangentes desde P a la circunferencia. Llamando, finalmente, (x, y) a las coordenadas genéricas de X, en lugar de (α, β) como se ha hecho hasta aquí, el lugar pedido es:

$$(x - x_0)^2[\rho^2 - (b - y_0)^2] + (y - y_0)^2[\rho^2 - (a - x_0)^2] +$$
$$+2(x - x_0)(y - y_0)(a - x_0)(b - y_0) = 0$$

Este lugar, que está formado por rectas (las tangentes desde P a la circunferencia), consta de dos rectas, ya que su ecuación es de segundo grado.

γ GEOMETRIA TRIDIMENSIONAL (AFIN Y EUCLIDEA)

Los problemas geométricos que nos van a ocupar a continuación son, esencialmente, los mismos que ya estudiamos en el caso bidimensional, si bien es cierto que, ahora, hay un mayor grado de dificultad y las situaciones son más variadas y complejas. A pesar de ello, no nos vamos a encontrar aquí con ninguna cuestión realmente difícil.

11.5. ECUACIONES DE RECTAS Y PLANOS

Recuérdese (véanse [202] y [203]) que las rectas y los planos son, respectivamente, los conjuntos de puntos del tipo:

$$\begin{array}{ll} r = P + \mathcal{V}(\bar{u}) & \pi = P + \mathcal{V}(\bar{u},\ \bar{v}) \\ r = P + \mathcal{V}(\overrightarrow{PQ}) \quad \text{y} & \pi = P + \mathcal{V}(\overrightarrow{PQ},\ \overrightarrow{PR}) \end{array}$$

donde, respectivamente: r es la recta que pasa por P y tiene la dirección de $\bar{u} \neq \bar{o}$, o que pasa por P y Q; y π es el plano que pasa por P y tiene la dirección que determinan $\bar{u}$ y $\bar{v}$ (independientes), o que pase por P, Q y R (no alineadas). Estas rectas y planos están formados por los puntos $X \in E_3$ que se pueden expresar, respectivamente, de las siguientes formas:

$$\begin{array}{ll} X = P + \lambda\bar{u} & X = P + \lambda\bar{u} + \mu\bar{v} \\ X = P + \lambda\overrightarrow{PQ} \quad (\lambda \in \mathbb{R}); & X = P + \lambda\overrightarrow{PQ} + \mu\overrightarrow{PR} \quad (\lambda, \mu \in \mathbb{R}) \end{array}$$

que son las ecuaciones vectoriales (paramétricas) de la recta r y del plano π.

No está de más insistir en que para determinar un punto de la recta r basta con asignar un valor al parámetro λ; para determinar un punto del plano π se precisan dos valores, uno para λ y otro para μ. Para señalar que las cosas ocurren de este modo, se dice que la recta y el plano son familias de puntos con uno y dos grados de libertad, respectivamente.

ECUACIONES DE LOS PLANOS

En el espacio E_3 se va a fijar una referencia cartesiana $(O;\ \bar{e}_1,\ \bar{e}_2,\ \bar{e}_3)$ y, con ello, cada punto $X \in E_3$ tendrá su sistema de coordenadas $(x_1,\ x_2,\ x_3)$, que le determinan de manera biunívoca. Dado un plano de E_3, vamos a obtener lo relación entre las coordenadas (x_i) que caracteriza a los puntos $X \in E_3$ que pertenecen al plano dado.

[215]

1. En el espacio tridimensional afín E_3, el plano de ecuación vectorial $X = P + \lambda\bar{u} + \mu\bar{v}$, que pasa por el punto $P(p_1, p_2, p_3)$ y tiene la dirección definida por los vectores independientes $\bar{u}(u_1, u_2, u_3)$ y $\bar{v}(v_1, v_2, v_3)$ de $\mathbb{R}^3$, admite las siguientes ecuaciones (*ecuaciones paramétricas*) con coordenadas:

$$\begin{cases} x_1 = p_1 + \lambda u_1 + \mu v_1 \\ x_2 = p_2 + \lambda u_2 + \mu v_2 \\ x_3 = p_3 + \lambda u_3 + \mu v_3 \end{cases} \quad (\text{para } \lambda, \mu \in \mathbb{R}) \qquad [1]$$

2. El plano que pasa por los tres puntos no alineados $P(p_1, p_2, p_3)$, $Q(q_1, q_2, q_3)$ y $R(r_1, r_2, r_3)$, de E_3, admite por ecuaciones a las:

$$\begin{cases} x_1 = p_1 + \lambda(q_1 - p_1) + \mu(r_1 - p_1) \\ x_2 = p_2 + \lambda(q_2 - p_2) + \mu(r_2 - p_2) \\ x_3 = p_3 + \lambda(q_3 - p_3) + \mu(r_3 - p_3) \end{cases} \quad (\text{para } \lambda, \mu \in \mathbb{R}) \quad \text{o}$$

$$\begin{vmatrix} x_1 & x_2 & x_3 & 1 \\ p_1 & p_2 & p_3 & 1 \\ q_1 & q_2 & q_3 & 1 \\ r_1 & r_2 & r_3 & 1 \end{vmatrix} = 0 \qquad [2]$$

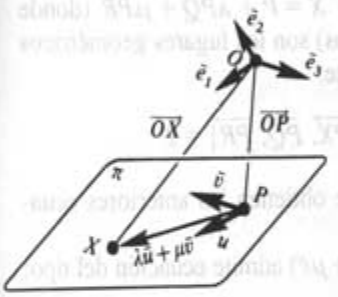


Figura 11.17.

3. Todo plano del espacio afín E_3 admite ecuación de la forma [3], que abajo se da, para ciertos números $a_0, a_1, a_2, a_3 \in \mathbb{R}$ con $(a_1, a_2, a_3) \neq \neq (0, 0, 0)$; recíprocamente, toda ecuación de dicho tipo representa a un plano de E_3. Si $P(p_1, p_2, p_3)$ es un punto de un plano, la ecuación general de éste puede ponerse en la forma [4] (ecuación general de un plano de E_3):

$$a_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0 \quad \text{o} \qquad [3]$$

$$a_1(x_1 - p_1) + a_2(x_2 - p_2) + a_3(x_3 - p_3) = 0 \qquad [4]$$

COMPROBACIONES

1. Los puntos del plano son aquellos $X(x_i)$ tales que $\overrightarrow{PX} = \lambda\bar{u} + \mu\bar{v}$ para algunos $\lambda, \mu \in \mathbb{R}$, o sea tales que $x_i - p_i = \lambda u_i + \mu v_i$ (para $i = 1, 2, 3$; con $\lambda, \mu \in \mathbb{R}$), que son las ecuaciones paramétricas del enunciado.
2. Repitiendo lo anterior para $\bar{u} = \overrightarrow{PQ}$ y $\bar{v} = \overrightarrow{PR}$, se obtienen las ecuaciones paramétricas [2]. Respecto de la ecuación «en forma de determinante», restando la segunda fila a cada una de las demás y desarrollando luego el determinante por su última columna, se ve que esta ecuación expresa que los vectores $\overrightarrow{PX}(x_i - p_i)$, $\overrightarrow{PQ}(q_i - p_i)$ y $\overrightarrow{PR}(r_i - p_i)$ son linealmente dependientes, esto es, que X pertenece al plano que pasa por P, Q y R.

3. De acuerdo con lo anterior, todo plano admite ecuación «en forma de determinante»; desarrollando éste por elementos de la primera fila, se obtiene la ecuación general, en la que los a_1, a_2 y a_3 que se obtienen no pueden ser los tres nulos, pues $P(p_i)$, $Q(q_i)$ y $R(r_i)$ son tres puntos no alineados (véanse las observaciones que se hacen a continuación). Recíprocamente, dada una ecuación del tipo [3] con $(a_1, a_2, a_3) \neq (0, 0, 0)$, por ejemplo, con $a_1 \neq 0$, es evidente que dicha ecuación es la del plano que pasa por el punto $P(-a_0/a_1, 0, 0)$ y tiene como vectores de su dirección a los $\bar{u}(-a_2, a_1, 0)$ y $\bar{v}(-a_3, 0, a_1)$, que son independientes.

OBSERVACIONES

- A las coordenadas de un punto genérico de E_3, en lugar de llamarlas x_1, x_2, x_3, se las suele denotar por x, y, z. Así, la ecuación general de un plano se escribirá poniendo $ax + by + cz + d = 0$, donde los números a, b, $c \in \mathbb{R}$ no pueden ser los tres nulos.
- Los planos de ecuaciones $X = P + \lambda\bar{u} + \mu\bar{v}$ y $X = P + \lambda\overrightarrow{PQ} + \mu\overrightarrow{PR}$ (donde $\bar{u}$ y $\bar{v}$ son independientes y P, Q y R no alineados) son los lugares geométricos de los puntos $X \in E_3$ tales que, respectivamente:

$$\text{rang } \{\overrightarrow{PX}, \bar{u}, \bar{v}\} = 2 \quad \text{y} \quad \text{rang } \{\overrightarrow{PX}, \overrightarrow{PQ}, \overrightarrow{PR}\} = 2$$

  Al expresar estas relaciones en coordenadas, se obtienen las anteriores ecuaciones [1] y [2] de un plano.
- Según el punto anterior, todo plano $(X = P + \lambda\bar{u} + \mu\bar{v})$ admite ecuación del tipo:

$$\begin{vmatrix} x_1 - p_1 & x_2 - p_2 & x_3 - p_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0 \quad \text{o}$$

$$\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} (x_1 - p_1) + \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix} (x_2 - p_2) + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} (x_3 - p_3) = 0$$

  Obsérvese que esta última es la ecuación general [4], cuyos coeficientes a_1, a_2, a_3 han resultado ser, pues, los menores de orden 2 de la matriz

$$\begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix}$$

  que no pueden ser los tres nulos, pues el rango de esta matriz es 2, ya que sus filas no son proporcionales, por no serlo $\bar{u}$ y $\bar{v}$.

EJERCICIO

Hallar la ecuación general del plano π', del espacio afín E_3, simétrico respecto de O, origen de coordenadas, del plano π que tiene por ecuación general a la $a_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0$.

RESOLUCION

La ecuación pedida es $\pi' : -a_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0$ ya que $X'(x_i)$ es de π' si y sólo si $X(-x_i)$ es de π, esto es, si y sólo si

$$a_0 + a_1(-x_1) + a_2(-x_2) + a_3(-x_3) = 0$$

que es la solución dada.

EXPRESION EUCLIDEA DE UN PLANO

[216]

Un plano π cualquiera, del espacio euclídeo E_3, queda definido si se conoce uno de sus puntos, P, y un vector no nulo $\bar{a} \in \mathbb{R}^3$ ortogonal a su dirección. En función de P y $\bar{a}$, el plano π es el lugar geométrico de los puntos $X \in E_3$ tales que:

$$\bar{a} \cdot \overrightarrow{PX} = 0 \quad \text{(ecuación euclídea del plano)} \qquad [1]$$

(si $\bar{u}$ y $\bar{v}$ son vectores de la dirección de π, independientes, tómese $\bar{a} = \bar{u} \wedge \bar{v}$).

Recurriendo a cualquier referencia cartesiana rectangular y si en ella $P(p_1, p_2, p_3)$ y $\bar{a}(a_1, a_2, a_3)$, el anterior plano π admite por ecuación a la:

$$a_1(x_1 - p_1) + a_2(x_2 - p_2) + a_3(x_3 - p_3) = 0 \qquad [2]$$

El vector $\bar{a}(a_1, a_2, a_3)$ es ortogonal (*) al plano $a_0 + a_1x + a_2x_2 + a_3x_3 = 0$.

(*) Abusando del lenguaje, un vector se dice ortogonal a un plano si lo es a su dirección.

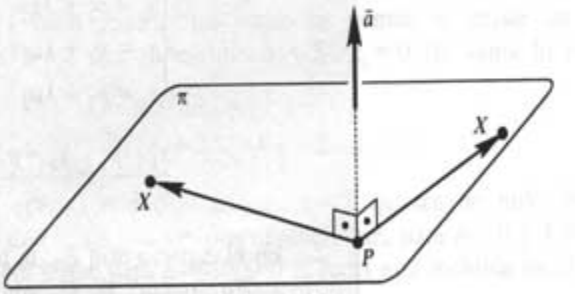


Figura 11.18.

DEMOSTRACION

Sean $\bar{u}$ y $\bar{v}$ dos de los muchos vectores de $\mathbb{R}^3$ ortogonales a $\bar{a}$ e independientes; es decir, $(\bar{u}, \bar{v})$ es una base del subespacio, de $\mathbb{R}^3$, ortogonal a $\bar{a}$. Decir que $\bar{a} \cdot \overrightarrow{PX} = 0$, esto es, que $\overrightarrow{PX}$ es ortogonal a $\bar{a}$, equivale pues a decir que $\overrightarrow{PX}$ es una combinación lineal de $\bar{u}$ y $\bar{v}$, o sea, que $X = P + \lambda\bar{u} + \mu\bar{v}$ para $\lambda, \mu \in \mathbb{R}$, lo que significa que X recorre un plano: el que pasa por P y tiene por dirección

al subespacio ortogonal a $\bar{a}$. Es obvio que el producto vectorial $\bar{u} \wedge \bar{v}$, por ser ortogonal a $\bar{u}$ y a $\bar{v}$, tiene la dirección del vector $\bar{a}$.

Utilizando una referencia rectangular cualquiera, en la que $\bar{a}(a_1, a_2, a_3)$ y $\overrightarrow{PX}(x_1 - p_1, x_2 - p_2, x_3 - p_3)$, la ecuación [1] toma la forma [2].

EJERCICIO

De una pirámide se conoce el vértice, A, y su proyección sobre la base, A'. Hallar la ecuación del plano de la base, siendo (en coordenadas rectangulares):

$$A(2, 1, -3) \quad \text{y} \quad A'(3, 4, 1)$$

RESOLUCION

La base está en el plano que pasa por el punto $A(2, 1, -3)$ y es ortogonal al vector $\overrightarrow{AA'}(1, 3, 4)$, luego su ecuación euclídea es:

$$1(x_1 - 2) + 3(x_2 - 1) + 4(x_3 + 3) = 0 \quad \text{o} \quad x_1 + 3x_2 + 4x_3 + 7 = 0$$

ECUACIONES DE LAS RECTAS (DE E_3)

Echando mano de una referencia cartesiana, en E_3, las rectas de este espacio se expresan como los lugares geométricos de los puntos $X \in E_3$ cuyas coordenadas (x_1, x_2, x_3) verifican a las siguientes relaciones:

[217]

1. En el espacio afín E_3, la recta cuya ecuación vectorial es $X = P + \lambda\bar{u}$, o sea, la que pasa por el punto $P(p_1, p_2, p_3)$, de E_3, y tiene la dirección del vector no nulo $\bar{u}(u_1, u_2, u_3)$, de $\mathbb{R}^3$, admite las siguientes ecuaciones con coordenadas:

$$\begin{cases} x_1 = p_1 + \lambda u_1 \\ x_2 = p_2 + \lambda u_2 \\ x_3 = p_3 + \lambda u_3 \end{cases} \quad \text{para } \lambda \in \mathbb{R} \quad \left(\begin{matrix}\text{ecuaciones} \\ \text{paramétricas}\end{matrix}\right) \text{o}$$

$$\frac{x_1 - p_1}{u_1} = \frac{x_2 - p_2}{u_2} = \frac{x_3 - p_3}{u_3} \quad \text{(ecuación continua)} \qquad [1]$$

2. En el espacio afín E_3, la recta que pasa por los puntos $P(p_1, p_2, p_3)$ y $Q(q_1, q_2, q_3)$, de E_3, admite por ecuaciones a las

$$\begin{cases} x_1 = p_1 + \lambda(q_1 - p_1) \\ x_2 = p_2 + \lambda(q_2 - p_2) \\ x_3 = p_3 + \lambda(q_3 - p_3) \end{cases} \quad \text{para } \lambda \in \mathbb{R} \quad \text{o}$$

$$\frac{x_1 - p_1}{q_1 - p_1} = \frac{x_2 - p_2}{q_2 - p_2} = \frac{x_3 - p_3}{q_3 - p_3} \qquad [2]$$

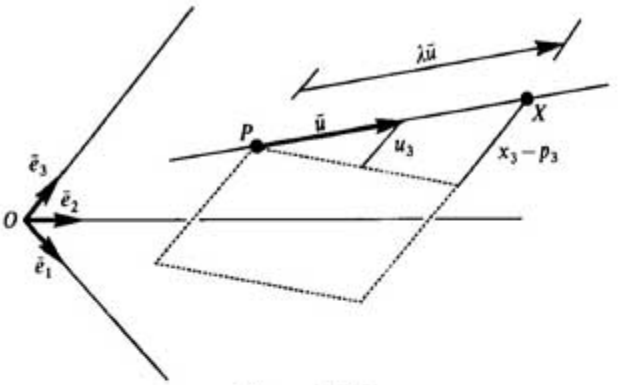


Figura 11.19.

COMPROBACION

1. Aquí vale lo que se dijo en la comprobación de [203], 1.
2. Basta con tomar $\bar{u} = \overrightarrow{PQ}$ en el apartado anterior.

EJERCICIO

Hallar la condición para que, en el espacio afín E_3, la recta r y el plano π se corten en un y sólo un punto:

$$r: \frac{x_1 - p_1}{u_1} = \frac{x_2 - p_2}{u_2} = \frac{x_3 - p_3}{u_3}; \qquad \pi: a_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0$$

RESOLUCION

El plano y la recta tendrán un punto en común si existe un $\lambda \in \mathbb{R}$ tal que las $x_i = p_i + \lambda u_i$ sean tales que $a_0 + \Sigma a_i x_i = 0$ (la suma lo es para $i = 1, 2, 3$), o sea, tal que

$$(a_0 + \Sigma a_i p_i) + \lambda \Sigma a_i u_i = 0$$

Así, pues, la condición pedida es que $h = \Sigma a_i u_i$ sea no nulo. Nótese que, llamando $k = a_0 + \Sigma a_i p_i$, se verifica también que, para $h = 0$ y $k \neq 0$ la intersección es vacía y que para $h = 0$ $k = 0$ la recta está incluida en el plano.

OBSERVACION

En una referencia cartesiana (coordenadas x, y, z), si la recta r que corta al plano $z = 0$ en el punto $(a, b, 0)$ y si (h, k, l) es un vector de su dirección, llamando $p = h/l$ y $q = k/l$, dicha recta admite por ecuación a:

$$r: \begin{cases} x = a + pz \\ y = b + qz \end{cases}$$

LAS RECTAS COMO INTERSECCIONES DE PLANOS

[218]

1. Toda recta del espacio afín E_3 es intersección de dos planos (distintos) y, por ello, está representada por un sistema de ecuaciones del tipo:

$$\begin{cases} a_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0 \\ b_0 + b_1x_1 + b_2x_2 + b_3x_3 = 0 \end{cases} \quad \text{donde} \quad \operatorname{rang} \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = 2 \quad [1]$$

2. En el espacio euclídeo E_3, de acuerdo con lo anterior, toda recta r se puede expresar como el lugar geométrico de los puntos $X \in E_3$ que verifican a un sistema de ecuaciones del tipo $\bar{a} \cdot \overrightarrow{PX} = 0$, $\bar{b} \cdot \overrightarrow{QX} = 0$, donde $P, Q \in E_3$ y $\bar{a}, \bar{b} \in \mathbb{R}^3$ son independientes. La recta r tiene la dirección del vector $\bar{u} = \bar{a} \wedge \bar{b}$.

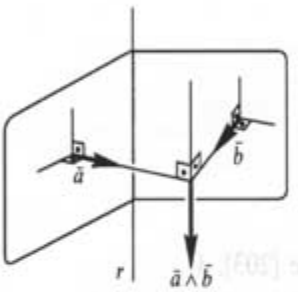


Figura 11.20.

COMPROBACIONES

1. Sea dada una recta r. Tomando un punto $P(p_1, p_2, p_3)$ de ella y un vector $\bar{u}(u_1, u_2, u_3)$ de su dirección (uno al menos de los u_i es no nulo; supondremos $u_1 \neq 0$), de acuerdo con [217], 1, la recta r está representada por

$$\begin{cases} u_2(x_1 - p_1) - u_1(x_2 - p_2) = 0 \\ u_3(x_1 - p_1) - u_1(x_3 - p_3) = 0 \end{cases} \quad \text{donde}$$

$$\operatorname{rang} \begin{bmatrix} u_2 & -u_1 & 0 \\ u_3 & 0 & -u_1 \end{bmatrix} = 2$$

que es, pues, del tipo [1] descrito en el enunciado.

Recíprocamente, sea r el conjunto de los puntos que satisfacen al sistema [1]. Según el teorema de Rouché, este sistema tiene infinitas soluciones, con un grado de libertad, de modo que si $(x_1, x_2, x_3) = (r_1, r_2, r_3)$ es una solución del sistema [1] y si $(x_1, x_2, x_3) = (u_1, u_2, u_3)$ es una solución del sistema homogéneo del [1], resulta que la solución general del sistema [1] es la

$$(x_1, x_2, x_3) = (r_1, r_2, r_3) + \lambda(u_1, u_2, u_3), \quad \text{para } \lambda \in \mathbb{R}$$

Por tanto, r es la recta que pasa por el punto $R(r_1, r_2, r_3)$ y tiene la dirección del vector $\bar{u}(u_1, u_2, u_3)$.

2. Las dos ecuaciones [1] lo son de planos respectivamente ortogonales a los vectores $\bar{a}(a_1, a_2, a_3)$ y $\bar{b}(b_1, b_2, b_3)$ (las coordenadas se suponen rectangulares); por tanto, recurriendo de un punto P del primero y a otro punto Q del segundo, los planos se pueden expresar en la forma $\bar{a} \cdot \overrightarrow{PX} = 0$ y $\bar{b} \cdot \overrightarrow{QX} = 0$. Hay que señalar que, como el rango de la matriz de [1] vale

2, se puede asegurar que $\bar{a}$ y $\bar{b}$ son vectores linealmente independientes. Para concluir, notemos que r, por pertenecer a ambos planos, tiene su dirección ortogonal a $\bar{a}$ y a $\bar{b}$ y, en consecuencia, dicha recta r tiene la dirección de $\bar{a} \wedge \bar{b}$.

EJERCICIO

Hallar un vector que tenga la dirección de la recta según la que se cortan los planos π_1 y π_2; el plano π_1 pasa por los puntos $A_1(0, 1, -2)$, $B_1(1, 2, 1)$ y $C_1(2, -2, -1)$ y el plano π_2 pasa por los puntos $A_2(1, -1, 2)$, $B_2(4, 1, 1)$ y $C_2(2, 1, 3)$ (las coordenadas son rectangulares).

RESOLUCION

Los siguientes vectores $\bar{a}_1$ y $\bar{a}_2$ son ortogonales a π_1 y π_2:

$$\bar{a}_1 = \overrightarrow{A_1B_1} \wedge \overrightarrow{A_1C_1} = (1, 1, 3) \wedge (2, -3, 1) = (10, 5, -5) = 5(2, -3, 1)$$

$$\bar{a}_2 = \overrightarrow{A_2B_2} \wedge \overrightarrow{A_2C_2} = (3, 2, -1) \wedge (1, 2, 1) = (4, -4, 4) = 4(1, -1, 1)$$

Por tanto, un vector $\bar{u}$ de la dirección de la recta $\pi_1 \cap \pi_2$ es el:

$$\bar{u} = (2, -3, 1) \wedge (1, -1, 1) = (-2, -1, 1)$$

11.6. POSICIONES RELATIVAS DE RECTAS Y PLANOS

Aun cuando la mayor parte de las cuestiones que aquí se van a analizar son de tipo afín, también hablaremos acerca de la ortogonalidad, que es materia euclídea. De este modo intentamos evitar las reiteraciones que se producirían si, de no hacer las cosas así, hubiéramos de tratar una misma cuestión, o distintos aspectos de ella, primero desde el punto de vista afín y, más tarde, bajo la óptica euclídea.

ORTOGONALIDAD

Aun cuando pueda resultar chocante, vamos a hablar antes de ortogonalidad que de paralelismo. Lo que ahora digamos va a sernos luego de utilidad, cuando hablemos de rectas y planos paralelos, y ello aun en el caso de que E_3 sea «solamente» espacio afín, pues se le puede dotar de estructura euclídea de un modo natural.

[219]

En el espacio euclídeo E_3, la ortogonalidad entre rectas y planos se define así:

1. Se dice que dos rectas son ortogonales (o perpendiculares) si lo son sus direcciones. Así, las rectas $P + V(\bar{u})$ y $Q + V(\bar{v})$ son ortogonales si $\bar{u} \cdot \bar{v} = 0$.
2. Se dice que una recta y un plano son ortogonales si lo son sus direcciones. Así, la recta $P + V(\bar{u})$ y el plano de ecuación euclídea $\bar{a} \cdot \overrightarrow{QX} = 0$ son ortogonales si $\bar{u}$ y $\bar{a}$ son proporcionales.
3. Se dice que dos planos π_1 y π_2 son ortogonales si lo son $\bar{a}_1$ y $\bar{a}_2$, donde $\bar{a}_1$ es ortogonal a π_1 y $\bar{a}_2$ lo es a π_2. Así, los planos de ecuaciones $\bar{a}_1 \cdot \overrightarrow{P_1X} = 0$ y $\bar{a}_2 \cdot \overrightarrow{P_2X} = 0$ son ortogonales si $\bar{a}_1 \cdot \bar{a}_2 = 0$.

COMPROBACIONES

En los tres casos, las condiciones de ortogonalidad (esto es: $\bar{u} \cdot \bar{v} = 0$; $\bar{u}$ y $\bar{a}$ son proporcionales; $\bar{a}_1 \cdot \bar{a}_2 = 0$) son consecuencias obvias de las correspondientes definiciones.

EJERCICIOS

Obtener condiciones de ortogonalidad para las rectas y los planos, del espacio euclídeo E_3, que tienen las siguientes ecuaciones vectoriales:

1. Recta $r: X = P + \lambda\bar{u}$ y recta $s: \bar{a}_1 \cdot \overrightarrow{P_1X} = 0$ y $\bar{a}_2 \cdot \overrightarrow{P_2X} = 0$.
2. Recta $r: X = P + \lambda\bar{u}$ y plano $\pi: X = Q + \lambda_1\bar{u}_1 + \lambda_2\bar{u}_2$.
3. Plano $\pi: \bar{a} \cdot \overrightarrow{PX} = 0$ y recta $r: \bar{a}_1 \cdot \overrightarrow{P_1X} = 0$ y $\bar{a}_2 \cdot \overrightarrow{P_2X} = 0$.
4. Plano $\pi: \bar{a} \cdot \overrightarrow{PX} = 0$ y plano $\tau: X = Q + \lambda\bar{u} + \mu\bar{v}$.

RESOLUCION

De la definición de ortogonalidad, se deduce trivialmente que:

1. $(r \perp s) \Leftrightarrow \bar{u} \cdot (\bar{a}_1 \wedge \bar{a}_2) = 0 \Leftrightarrow \operatorname{rang}(\bar{u}, \bar{a}_1, \bar{a}_2) = 2$.
2. $(r \perp \pi) \Leftrightarrow \bar{u} \cdot \bar{u}_1 = \bar{u} \cdot \bar{u}_2 = 0 \Leftrightarrow \bar{u}$ y $\bar{u}_1 \wedge \bar{u}_2$ proporcionales.
3. $(\pi \perp r) \Leftrightarrow \bar{a} \cdot \bar{a}_1 = \bar{a} \cdot \bar{a}_2 = 0 \Leftrightarrow \bar{a}$ y $\bar{a}_1 \wedge \bar{a}_2$ proporcionales.
4. $(\pi \perp \tau) \Leftrightarrow \bar{a} \cdot (\bar{u} \wedge \bar{v}) = 0 \Leftrightarrow \operatorname{rang}(\bar{a}, \bar{u}, \bar{v}) = 2$.

PARALELISMO

[220]

En el espacio afín E_3, el paralelismo, entre rectas y planos, se define como sigue:

1. Dos rectas se dicen paralelas si tienen la misma dirección. Así, las rectas $P + V(\bar{u})$ y $Q + V(\bar{v})$ son paralelas si $\bar{u}$ y $\bar{v}$ son proporcionales.
2. Dos planos se dicen paralelos si tienen la misma dirección. Así, los planos de ecuaciones euclídeas (*) $\bar{a} \cdot \overrightarrow{PX} = 0$ y $\bar{b} \cdot \overrightarrow{QX} = 0$ son paralelos si $\bar{a}$ y $\bar{b}$ son proporcionales.
3. Una recta se dice paralela a un plano si la dirección de la recta está incluida en la del plano. Así, la recta $P + V(\bar{u})$ y el plano de ecuación euclídea (*) $\bar{a} \cdot \overrightarrow{QX} = 0$ son paralelos si $\bar{u} \cdot \bar{a} = 0$.

(*) Si el espacio afín E_3 no fuese euclídeo, estos planos se deberían poner en las formas $a_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0$ y $b_0 + b_1x_1 + b_2x_2 + b_3x_3 = 0$ y llamar $\bar{a} = (a_1, a_2, a_3)$ y $\bar{b} = (b_1, b_2, b_3)$. La relación $\bar{u} \cdot \bar{a} = 0$ será la $u_1a_1 + u_2a_2 + u_3a_3 = 0$. Nótese que al espacio afín E_3 se le da estructura euclídea si se considera el producto escalar canónico, o sea, con $(u_1, u_2, u_3) \cdot (v_1, v_2, v_3) = u_1v_1 + u_2v_2 + u_3v_3$.

COMPROBACIONES

En los tres casos, las condiciones de paralelismo que se dan son consecuencia inmediata de las definiciones que las preceden.

EJERCICIO

Obtener condiciones de paralelismo para las rectas y los planos, de E_3 (afín o afín euclídeo, si fuera preciso), que tienen las siguientes ecuaciones vectoriales:

1. Recta $r: X = P + \lambda\bar{u}$ y recta $s: \bar{a}_1 \cdot \overrightarrow{P_1X} = 0$ y $\bar{a}_2 \cdot \overrightarrow{P_2X} = 0$.

2. Recta $r: a_1 \cdot \overrightarrow{P_1X} = 0$ y $\bar{a}_2 \cdot \overrightarrow{P_2X} = 0$ y recta $s: \bar{b}_1 \cdot \overrightarrow{Q_1X} = 0$ y $\bar{b}_2 \cdot \overrightarrow{Q_2X} = 0$.

3. Recta $r: X = P + \lambda\bar{u}$ y plano $\pi: X = Q + \lambda_1\bar{u}_1 + \lambda_2\bar{u}_2$.

4. Recta $r: \bar{a}_1 \cdot \overrightarrow{P_1X} = 0$ y $\bar{a}_2 \cdot \overrightarrow{P_2X} = 0$ y plano $\pi: \bar{b} \cdot \overrightarrow{QX} = 0$.

5. Plano $\pi: X = P + \lambda\bar{u} + \mu\bar{v}$ y plano $\tau: \bar{a} \cdot \overrightarrow{QX} = 0$.

6. Plano $\pi: X = P + \lambda_1\bar{u}_1 + \lambda_2\bar{u}_2$ y plano $\tau: X = Q + \mu_1\bar{v}_1 + \mu_2\bar{v}_2$.

RESOLUCION

De la definición de paralelismo se deduce trivialmente que (el símbolo // indica «paralelo a»):

1. $(r//s) \Leftrightarrow \bar{u}$ y $\bar{a}_1 \wedge \bar{a}_2$ proporcionales $\Leftrightarrow \bar{u} \cdot \bar{a}_1 = \bar{u} \cdot \bar{a}_2 = 0$.

2. $(r//s) \Leftrightarrow \text{rang}\,(\bar{a}_1, \bar{a}_2, \bar{b}_1, \bar{b}_2) = 2 \Leftrightarrow \bar{a}_1 \wedge \bar{a}_2$ y $\bar{b}_1 \wedge \bar{b}_2$ proporcionales.
3. $(r//\pi) \Leftrightarrow \text{rang}\,(\bar{u}, \bar{u}_1, \bar{u}_2) = 2 \Leftrightarrow \bar{u} \cdot (\bar{u}_1 \wedge \bar{u}_2) = 0.$
4. $(r//\pi) \Leftrightarrow \text{rang}\,(\bar{a}_1, \bar{a}_2, \bar{b}) = 2 \Leftrightarrow (\bar{a}_1 \wedge \bar{a}_2) \cdot \bar{b} = 0.$
5. $(\pi//\tau) \Leftrightarrow \bar{u} \wedge \bar{v}$ y $\bar{a}$ proporcionales $\Leftrightarrow \bar{u} \cdot \bar{a} = \bar{v} \cdot \bar{a} = 0.$
6. $(\pi//\tau) \Leftrightarrow \text{rang}\,(\bar{u}_1, \bar{u}_2, \bar{v}_1, \bar{v}_2) = 2 \Leftrightarrow \bar{u}_1 \wedge \bar{u}_2$ y $\bar{v}_1 \wedge \bar{v}_2$ proporcionales.

INTERSECCIONES DE PLANOS

[221]

En el espacio E_3, y respecto de una referencia cartesiana de él, se verifica:

INTERSECCION DE DOS PLANOS: Sean dados los siguientes planos π y τ; llamemos M y M' a las matrices que abajo se señalan:

$$\pi: a_1x_1 + a_2x_2 + a_3x_3 + a_0 = 0$$
$$\tau: b_1x_1 + b_2x_2 + b_3x_3 + b_0 = 0$$

$$M = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \; ; \; M' = \begin{bmatrix} a_1 & a_2 & a_3 & a_0 \\ b_1 & b_2 & b_3 & b_0 \end{bmatrix}$$

- Si rang $M = 2$, entonces $\pi \cap \tau$ es una recta (π y τ se dicen secantes).
- Si rang $M = 1$ y rang $M' = 2$, entonces $\pi \cap \tau = \emptyset$ y π y τ son paralelos.
- Si rang M = rang $M' = 1$, entonces los planos coinciden (son el mismo).

INTERSECCION DE TRES PLANOS: Sean dados los siguientes planos π, τ y σ; llamemos M y M' a las matrices que abajo se señalan:

$$\pi: a_1x_1 + a_2x_2 + a_3x_3 + a_0 = 0$$
$$\tau: b_1x_1 + b_2x_2 + b_3x_3 + b_0 = 0$$
$$\sigma: c_1x_1 + c_2x_2 + c_3x_3 + c_0 = 0$$

$$M = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \; ; \; M' = \begin{bmatrix} a_1 & a_2 & a_3 & a_0 \\ b_1 & b_2 & b_3 & b_0 \\ c_1 & c_2 & c_3 & c_0 \end{bmatrix}$$

- Si rang $M = 3$, entonces $\pi \cap \tau \cap \sigma$ consta de un solo punto (forman un triedro).
- Si rang $M = 2$ y rang $M' = 3$, entonces $\pi \cap \tau \cap \sigma = \emptyset$ y los tres no son paralelos.
- Si rang M = rang $M' = 2$, entonces $\pi \cap \tau \cap \sigma$ es una recta.
- Si rang $M = 1$ y rang $M' = 2$, entonces $\pi \cap \tau \cap \sigma = \emptyset$ y los tres son paralelos.
- Si rang M = rang $M' = 1$, entonces π, τ y σ coinciden (son el mismo)

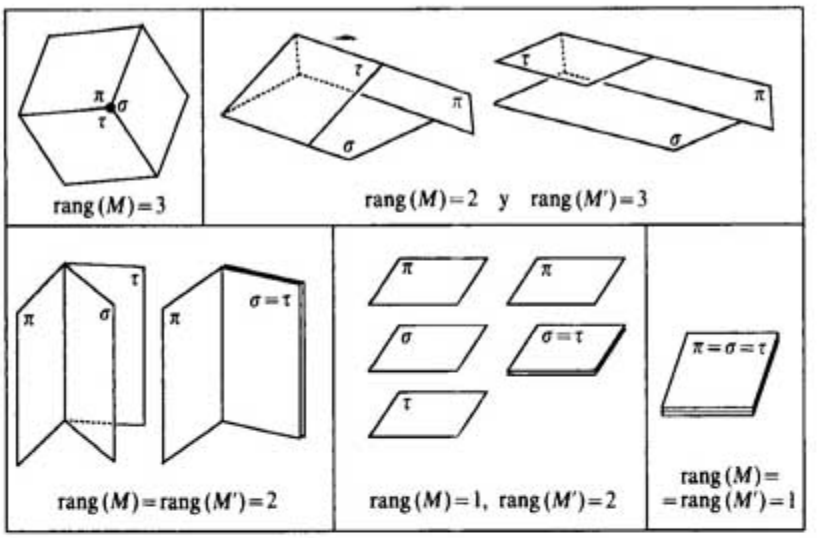


Figura 11.21.

COMPROBACIONES

Un punto $X(x_1, x_2, x_3)$ pertenecerá a la intersección de los planos si, y sólo si, (x_1, x_2, x_3) es solución del sistema que forman las ecuaciones de dichos planos. Al recurrir al teorema de Rouché, relativo a la discusión de sistemas de ecuaciones lineales, se obtienen los resultados que se dicen en el enunciado. En los casos en los que la intersección es una recta, que son las únicas no triviales, se puede razonar como se hizo en [218].

EJERCICIO

Estudiar las posiciones relativas de los tres planos siguientes, en función de los valores que tomen los parámetros $\alpha, \beta \in \mathbb{R}$:

$$x - y + 3z = 1 \quad ; \quad 3x - 5y + 7z + \beta = 0 \quad ; \quad x - 3y + \alpha z = 2$$

RESOLUCION

Recurriendo a las matrices M y M', del sistema que forman las tres ecuaciones dadas, y a sus rangos, se obtiene:

$$M = \begin{bmatrix} 1 & -1 & 3 \\ 3 & -5 & 7 \\ 1 & -3 & \alpha \end{bmatrix} \quad ; \quad M' = \begin{bmatrix} 1 & -1 & 3 & 1 \\ 3 & -5 & 1 & -\beta \\ 1 & -3 & \alpha & 2 \end{bmatrix}$$

	$\alpha \neq 1$ $\forall \beta$	$\alpha = 1$ $\beta \neq -4$	$\alpha = 1$ $\beta = -4$
rang M	3	2	2
rang M'	3	3	2

- Si $\alpha \neq 1$, los tres planos se cortan en un punto (forman un triedro).
- Si $\alpha = 1$ y $\beta \neq -4$, los tres planos tienen intersección vacía y forman un prisma (de aristas paralelas al vector $(4, 1, -1)$).
- Si $\alpha = 1$ y $\beta = -4$, los tres planos se cortan a lo largo de una recta, que es la $(x, y, z) = (1/2, -1/2, 0) + \lambda(4, 1, -1)$.

HACES DE PLANOS

[222]

En el espacio afín E_3, y respecto de una cierta referencia cartesiana, es:

Dados dos planos distintos cualesquiera $\pi: a_1x_1 + a_2x_2 + a_3x_3 + a_0 = 0$ y $\tau: b_1x_1 + b_2x_2 + b_3x_3 + b_0 = 0$, se llama *haz* de planos, que determinan π y τ, al conjunto de todos los planos que tienen ecuación del siguiente tipo:

$$\alpha(a_1x_1 + a_2x_2 + a_3x_3 + a_0) + \beta(b_1x_1 + b_2x_2 + b_3x_3 + b_0) = 0$$

para cualesquiera $\alpha, \beta \in \mathbb{R}$ tales que $(\alpha, \beta) \neq (0, 0)$ o

$$(a_1x_1 + a_2x_2 + a_3x_3 + a_0) + \lambda(b_1x_1 + b_2x_2 + b_3x_3 + b_0) = 0$$

para $\lambda \in \mathbb{R}^{(*)}$. Este haz está formado por los siguientes planos:

1. Si π y τ se cortan, en una recta r, entonces el haz lo forman todos los planos que pasan por r.
2. Si π y τ son paralelos, el haz lo forman todos los planos paralelos a los π y τ.

(*) Tomese $\lambda = \beta/\alpha$. Esta forma de dar el haz es defectuosa, pues no incluye al plano τ, que se obtiene para $(\alpha, \beta) = (0, 1)$. Para intentar remediarlo, se admite que τ es el límite del plano genérico del haz cuando $\lambda \to \infty$.

COMPROBACION

Para simplificar la notación, llamaremos $a(x)$ y $b(x)$ a:

$$a(x) = a_1x_1 + a_2x_2 + a_3x_3 + a_0 \quad \text{y} \quad b(x) = b_1x_1 + b_2x_2 + b_3x_3 + b_0$$

Nótese que por cualquier punto $P \in E_3$, excepto por los posibles puntos de la intersección $\pi \cap \tau$, pasa un y sólo un plano del haz, que es el que se obtiene cuando se toma $\alpha = b(p)$ y $\beta = -a(p)$.

1. *Si* π y τ se cortan en una recta r, entonces todos los planos del haz pasan por r, ya que para cualquiera que sea el punto $Q \in r$, como $a(q) = 0$ y $b(q) = 0$, sera $\alpha a(q) + \beta b(q) = 0$ para todos los $\alpha, \beta \in \mathbb{R}$. Recíprocamente, si σ es un plano que pasa por r, entonces σ ha de ser uno de los planos del haz, ya que, tomando un punto P de el que no esté en r, según se dijo antes, por P pasa un único plano σ' del haz y, como σ y σ' pasan por r y por P, ha de ser $\sigma' = \sigma$, luego σ es del haz.
2. Si π y τ son paralelos, esto es, si $\bar{a} = (a_1, a_2, a_3)$ y $\bar{b} = (b_1, b_2, b_3)$ son proporcionales, entonces todos los planos del haz son paralelos a ellos, pues $(\alpha a_1 + \beta b_1,\ \alpha a_2 + \beta b_2,\ \alpha a_3 + \beta b_3)$ es ahora proporcional a los $\bar{a}$ y $\bar{b}$. Recíprocamente, si σ es un plano paralelo a los π y τ, entonces σ ha de ser del haz, ya que, tomando un punto P cualquiera de σ, según se dijo al principio, por P pase un solo plano σ' del haz y, como σ y σ' son paralelos entre sí y pasan, ambos, por P, ha de ser $\sigma' = \sigma$, luego σ es del haz.

EJERCICIO (RADIACION DE PLANOS)

Sea dado un punto P de E_3. Se llama radiación de planos determinada por P al conjunto de todos los planos que pasan por P. Respecto de una referencia cartesiana cualquiera de E_3, se pide:

1. Si las coordenadas de P son (p_1, p_2, p_3), hallar la ecuación general de los planos de la radiación.

2. Si P es el punto de intersección de tres planos π, τ y σ (que forman triedro) y si las ecuaciones de éstos son:

$$\pi: \Sigma a_i x_i + a_0 = 0 \quad ; \quad \tau: \Sigma b_i x_i + b_0 = 0 \quad ; \quad \sigma: \Sigma c_i x_i + c_0 = 0$$

hallar la ecuación general de los planos de la radiación.

RESOLUCION

1. Se trata de la ecuación, genérica, de un plano cualquiera que pase por $P(p_1, p_2, p_3)$, que (en función de un vector (u_1, u_2, u_3) ortogonal al plano) es, según sabemos:

$$u_1(x_1 - p_1) + u_2(x_2 - p_2) + u_3(x_3 - p_3) = 0, \quad \text{para } (u_1, u_2, u_3) \in \mathbb{R}^3 - O$$

2. Los planos de la radiación son los:

$$\alpha(\Sigma a_i x_i + a_0) + \beta(\Sigma b_i x_i + b_0) + \gamma(\Sigma c_i x_i + c_0) = 0, \quad (\alpha, \beta, \gamma) \in \mathbb{R}^3 - O \,[1]$$

Es obvio que todos estos planos pasan por P, pues para él se anulan los tres paréntesis de la expresión [1]. También se verifica que cualquier plano

que pase por P es uno de los dados por [1]. Esto se comprueba tomando dos puntos Q y R en él y viendo que se pueden hallar α, β y γ para que [1] se verifique para Q y R, lo que no tiene dificultad.

INTERSECCION DE UNA RECTA Y UN PLANO

Según que a la recta se la conozca mediante ecuaciones generales (intersección de dos planos) o mediante ecuaciones paramétricas, se realizará un estudio distinto del problema (ambos conducirán, obviamente, al mismo resultado):

[223]

En el espacio afín y respecto de una cierta referencia cartesiana, se consideran la recta r y el plano π de ecuaciones generales:

$$r\begin{cases}\tau_1: a_1x_1 + a_2x_2 + a_3x_3 + a_0 = 0\\ \tau_2: b_1x_1 + b_2x_2 + b_3x_3 + b_0 = 0\end{cases} \quad \text{con} \quad \operatorname{rang}\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{bmatrix} = 2$$

$$\pi: c_1x_1 + c_2x_2 + c_3x_3 + c_0 = 0, \quad \text{siendo} \quad (c_1, c_2, c_3) \neq (0, 0, 0)$$

La intersección $r \cap \pi$ es, en función de los rangos de las matrices M y M', que se dan más abajo:

- Si rang $M = 3$, entonces r y π se cortan en un punto (único).
- Si rang $M = 2$ y rang $M' = 3$, entonces $r \cap \pi = \emptyset$ y r es paralela a π.
- Si rang M = rang $M' = 2$, entonces r está incluida en π.

$$M = \begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix} \quad M' = \begin{bmatrix}a_1 & a_2 & a_3 & a_0\\ b_1 & b_2 & b_3 & b_0\\ c_1 & c_2 & c_3 & c_0\end{bmatrix}$$

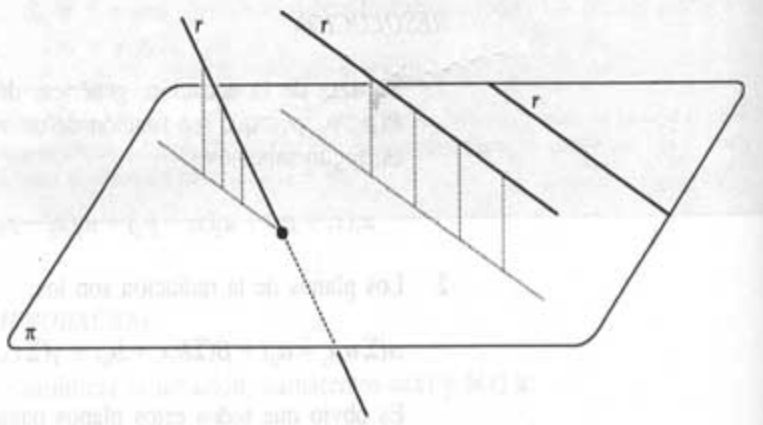


Figura 11.22.

COMPROBACION

Los resultados anteriores son parte de los ya considerados en [221], al estudiar la intersección de tres planos. Como ahora τ_1 y τ_2 son dos planos que se cortan (en r), ha de ser rang $M \geqslant 2$. Con esta limitación, lo dicho en [221] nos conduce fácilmente a estos resultados de ahora.

[224]

En el espacio afín E_3 y respecto de cierta referencia cartesiana, se da la recta r (en paramétricas) y el plano π (por su ecuación general):

$$r: \begin{cases} x_1 = p_1 + \lambda u_1 \\ x_2 = p_2 + \lambda u_2 \\ x_3 = p_3 + \lambda u_3 \end{cases} \quad (\text{para } \lambda \in \mathbb{R})$$

$$\pi: a_1x_1 + a_2x_2 + a_3x_3 + a_0 = 0$$

(nótese que $\vec{u} = (u_1, u_2, u_3)$ y $\vec{a} = (a_1, a_2, a_3)$ son no nulos).
La intersección $r \cap \pi$ es, en función de los siguientes valores de h y k:

$$h = a_1u_1 + a_2u_2 + a_3u_3 \quad \text{y} \quad k = a_1p_1 + a_2p_2 + a_3p_3 + a_0$$

- Si $h \neq 0$, entonces r y π se cortan en un punto (uno sólo).
- Si $h = 0$ y $k \neq 0$, entonces $r \cap \pi = \emptyset$ y r es paralela a π.
- Si $h = 0$ y $k = 0$, entonces r está incluida en π.

COMPROBACION

Estos resultados se obtuvieron ya anteriormente (véase el Ejercicio de [217]).

EJERCICIO

Estudiar la posición relativa de la recta r y el plano π siguientes:

$$r: \frac{x+1}{5} = \frac{y-2}{3} = \frac{z-4}{-2} \quad ; \quad \pi: \alpha x + y - z + \beta = 0$$

RESOLUCION

Para que un punto de la recta $(-1 + 5\lambda, 2 + 3\lambda, 4 - 2\lambda)$, esté en el plano, ha de ser

$$\alpha(-1 + 5\lambda) + (2 + 3\lambda) - (4 - 2\lambda) + \beta = 0 \quad \text{o sea} \quad 5(1 + \alpha)\lambda = 2 + \alpha - \beta$$

Por tanto, será: si $\alpha \neq -1$, entonces r y π se cortan en un punto; si $\alpha = -1$ y $\beta \neq 1$, entonces r es paralela a π; y si $\alpha = -1$ y $\beta = 1$, entonces r está incluida en π.

POSICION RELATIVA DE DOS RECTAS

Según que las rectas se conozcan mediante ecuaciones paramétricas o a través de ecuaciones generales (intersecciones de planos), se realizará el correspondiente estudio:

[225]

Dadas las rectas $r = P + V(\bar{u})$ y $s = Q + V(\bar{v})$, del espacio afín E_3, se verifica que:

- $[\text{rang}\,(\bar{u}, \bar{v}, \overrightarrow{PQ}) = 3] \Rightarrow r$ y s se cruzan (no se cortan y no son paralelas).
- $[\text{rang}\,(\bar{u}, \bar{v}) = \text{rang}\,(\bar{u}, \bar{v}, \overrightarrow{PQ}) = 2] \Rightarrow r$ y s se cortan en un punto (único).
- $[\text{rang}\,(\bar{u}, \bar{v}) = 1$ y $\text{rang}\,(\bar{u}, \bar{v}, \overrightarrow{PQ}) = 2] \Rightarrow r \cap s = \emptyset$ y r y s son paralelas.
- $[\text{rang}\,(\bar{u}, \bar{v}, \overrightarrow{PQ}) = 1] \Rightarrow r$ y s coinciden (son la misma recta).

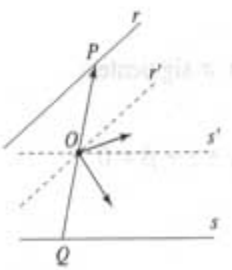
COMPROBACION

Para verificar los anteriores asertos, basta con tener en cuenta dos cosas. La primera de ellas es que el paralelismo de r y s queda caracterizado por la relación $\text{rang}\,(\bar{u}, \bar{v}) = 1$. La segunda es que r y s tendrán algún punto en común si, y sólo si, existen algunos $\lambda, \mu \in \mathbb{R}$ tales que $P + \lambda\bar{u} = Q + \mu\bar{v}$, esto es $\overrightarrow{PQ} = \lambda\bar{u} - \mu\bar{v}$, con lo que $\overrightarrow{PQ}$ depende linealmente de los $\bar{u}$ y $\bar{v}$.

EJERCICIO

Sean dadas dos rectas r y s, que se cruzan en el espacio. Describir cómo se determinaría una referencia cartesiana (coordenadas x, y, z) en la que r y s admitan las siguientes ecuaciones:

$$r: \begin{cases} x + y = 0 \\ z = 1 \end{cases} \qquad s: \begin{cases} x - y = 0 \\ z = -1 \end{cases}$$



RESOLUCION

Sea $(O; \bar{e}_1, \bar{e}_2, \bar{e}_3)$ la referencia buscada. Tómese un punto $P \in r$ y otro punto $Q \in s$; el origen O de la referencia puede ser el punto medio del segmento PQ. El vector $\bar{e}_3$ será el $\bar{e}_3 = \overrightarrow{OP}$.

Si $\bar{u}$ es un vector de la dirección de r y si $\bar{v}$ es un vector de la dirección de s, entonces se puede tomar $\bar{e}_1 = \bar{u} + \bar{v}$ y $\bar{e}_2 = \bar{u} - \bar{v}$.

[226]

En el espacio afín E_3 y respecto de una cierta referencia cartesiana, se consideran las rectas r y s de ecuaciones generales:

$$r:\begin{cases} a_1x_1+a_2x_2+a_3x_3+a_0=0 \\ b_1x_1+b_2x_2+b_3x_3+b_0=0 \end{cases} \text{ siendo } \operatorname{rang}\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}=2$$

$$s:\begin{cases} \alpha_1x_1+\alpha_2x_2+\alpha_3x_3+\alpha_0=0 \\ \beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_0=0 \end{cases} \text{ siendo } \operatorname{rang}\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{bmatrix}=2$$

La posición relativa de r y s es, dependiendo de los rangos de las matrices M y M' que se dan más abajo:

- Si rang $M'=4$, entonces r y s se cruzan (ni se cortan ni son paralelas).
- Si rang $M=$ rang $M'=3$, entonces r y s se cortan en un punto (único).
- Si rang $M=2$ y rang $M'=3$, entonces $r\cap s=\varnothing$ y r y s son paralelas.
- Si rang $M=$ rang $M'=2$, entonces r y s coinciden (son la misma recta).

$$M=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{bmatrix} \qquad M'=\begin{bmatrix} a_1 & a_2 & a_3 & a_0 \\ b_1 & b_2 & b_3 & b_0 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_0 \\ \beta_1 & \beta_2 & \beta_3 & \beta_0 \end{bmatrix}$$

COMPROBACION

Basta con aplicar el teorema de Rouché a la discusión del sistema que forman las ecuaciones de r y de s. También hay que recordar que el paralelismo de r y s (véase el Ejercicio de [220]) significa que el sistema que forman los cuatro vectores $\bar{a}=(a_i)$, $\bar{b}=(b_i)$, $\bar{\alpha}=(\alpha_i)$ y $\bar{\beta}=(\beta_i)$ tiene rango 2.

EJERCICIO

Estudiar (en función de α y β) la posición relativa de las rectas r y s:

$$r:(x,\ y,\ z)=(1,\ \alpha,\ 2)+\lambda(1,\ 1,\ \beta)\quad (\lambda\in\mathbb{R})\quad ;\quad s:x=y=z$$

RESOLUCION

Llevando a las ecuaciones de s los valores $x=1+\lambda$, $y=\alpha+\lambda$ y $z=2+\beta\lambda$, se obtiene:

$$1+\lambda=\alpha+\lambda=2+\beta\lambda$$

luego: si $\alpha=1$ y $\beta\neq 1$, las rectas se cortan; si $\beta=1$, entonces las rectas son paralelas; en los demás casos, las rectas se cruzan.

11.7. ANGULOS Y DISTANCIAS (EN E_3)

Hasta ahora, salvo lo dicho acerca de la ortogonalidad (véase [219]), sólo se han tratado cuestiones de tipo afín. De aquí en adelante, supondremos que el espacio afín E_3 es euclídeo, esto es, que disponemos del producto escalar entre vectores libres (vectores de $\mathbb{R}^3$). Cuando, de aquí en adelante, precisemos de una referencia cartesiana, tomaremos una que sea rectangular y, así, se abreviarán de modo apreciable las expresiones de los ángulos y de las distancias. La distancia de un punto $X(x_1, x_2, x_3)$ a otro punto $Y(y_1, y_2, y_3)$, expresadas en coordenadas rectangulares, vendrá dada por

$$d(X, Y) = \|\overrightarrow{XY}\| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2}$$

ANGULO DE DOS RECTAS

Aquí, en E_3, el concepto de ángulo de dos rectas es idéntico al que ya dimos anteriormente (véase [209]) para rectas de E_2:

[227]

Se llama *ángulo* ang (r, s), de dos rectas $r = P + V(\vec{u})$ y $s = Q + V(\vec{v})$, del espacio euclídeo E_3, al ángulo comprendido entre 0 y $\pi/2$ de entre los cuatro ángulos que forma un vector de los $\pm\vec{u}$ con otro de los $\pm\vec{v}$(*). El ang (r, s) queda determinado por su coseno, que vale:

$$\cos(r, s) = |\cos(\vec{u}, \vec{v})|$$

(*) Si a este ángulo se le llama α, los cuatro ángulos de $\pm\vec{u}$ y $\pm\vec{v}$ son los $\pm\alpha$ y $\pm(\pi - \alpha)$. Este ángulo α no varía si se sustituyen $\vec{u}$ y $\vec{v}$ por $\lambda\vec{u}$ y $\mu\vec{v}$ ($\lambda \neq 0$ y $\mu \neq 0$ cualesquiera).

COMPROBACION

Aquí es de aplicación lo dicho en los dos primeros párrafos de la comprobación de [209].

EJERCICIO

Hallar el ángulo que forman las siguientes rectas r y s (en coordenadas rectangulares):

$$r: \frac{x-3}{2} = \frac{y+5}{3} = z - 4; \qquad s: \begin{cases} 2x - z + 6 = 0 \\ 4x + 6y + z = 0 \end{cases}$$

RESOLUCION

Recurriendo a dos vectores $\bar{u}$ y $\bar{v}$ de las direcciones de r y s, se tiene:

$$\bar{u} = (2, 3, 1); \quad \bar{v} = (2, 0, -1) \wedge (4, 6, 1) = (6, -6, 12)$$

$$\cos(r, s) = |\cos(\bar{u}, \bar{v})| = \frac{|12 - 18 + 12|}{\sqrt{14}\sqrt{216}} \simeq 0{,}109$$

luego áng $(r, s) \simeq 83{,}7°$.

ANGULOS DE DOS PLANOS Y DE RECTA Y PLANO

Recurriendo al concepto de ángulo de dos rectas, se definen sin dificultad los ángulos entre dos planos y ángulo que forman una recta y un plano:

[228]

1. Se llama ángulo áng(π_1, π_2), de dos planos π_1 y π_2, del espacio euclídeo E_3, al ángulo que forma una recta perpendicular a π_1 con otra perpendicular a π_2. Si las ecuaciones euclídeas de los planos son $\pi_1 : \bar{a}_1 \cdot \overrightarrow{P_1X} = 0$ y $\pi_2 : \bar{a}_2 \cdot \overrightarrow{P_2X} = 0$, el coseno del áng (π_1, π_2) vale $\cos(\pi_1, \pi_2) = |\cos(\bar{a}_1, \bar{a}_2)|$.

2. Se llama ángulo áng (r, π), de la recta r y del plano π, del espacio euclídeo E_3, al ángulo que forma r con su proyección ortogonal sobre π, que es el complementario del ángulo que forma r con una recta perpendicular a π. Si r es la recta $P + V(\bar{u})$ y el plano π tiene por ecuación euclídea a $\bar{a} \cdot \overrightarrow{QX} = 0$, el seno del ángulo áng (r, π), vale $\operatorname{sen}(r, \pi) = |\cos(\bar{u}, \bar{a})|$.

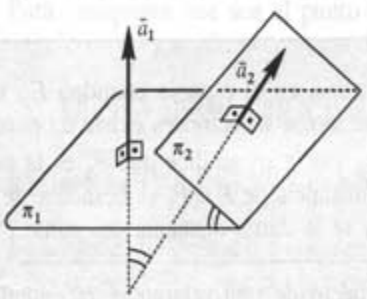


Figura 11.23.

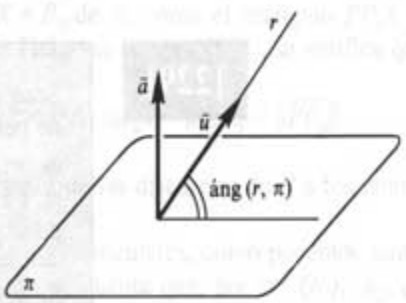


Figura 11.24.

COMPROBACION

A la vista de las definiciones anteriores, las expresiones que arriba se dan para $\cos(\pi_1, \pi_2)$ y para $\operatorname{sen}(r, \pi)$ se obtienen trivialmente.

EJERCICIOS

1. Dados dos planos π_1 y π_2, del espacio euclídeo E_3, compruébese que el áng (π_1, π_2) es el mayor de los ángulos de una recta de π_1 y una recta de π_2 que formen ángulos iguales con la recta $\pi_1 \cap \pi_2$.
2. Dados una recta r y un plano π, del espacio euclídeo E_3, compruébese que el áng (r, π) es el menor de los ángulos que forman r con una recta de π.

RESOLUCION

1. Sean r_1 y r_2 las intersecciones de π_1 y π_2 con un plano π perpendicular a la recta $r = \pi_1 \cap \pi_2$; obviamente, áng (π_1, π_2) = áng (r_1, r_2). Sean $\bar{u}$, $\bar{u}_1$ y $\bar{u}_2$ vectores unitarios de las direcciones de r, r_1 y r_2 (nótese que $\bar{u} \perp \bar{u}_1$ y que $\bar{u} \perp \bar{u}_2$). Hay que comprobar que $|\cos(\bar{u}_1, \bar{u}_2)|$ es el menor de todos los $|\cos(\bar{u}_1 + a\bar{u}, \bar{u}_2 + a\bar{u})|$ para $a \in \mathbb{R}$. Así es, ya que

$$\cos(\bar{u}_1 + a\bar{u}, \bar{u}_2 + a\bar{u}) = \frac{\bar{u}_1 \cdot \bar{u}_2 + a^2 \bar{u} \cdot \bar{u}}{\sqrt{1+a^2}\sqrt{1+a^2}} = \frac{\cos(\bar{u}_1, \bar{u}_2) + a^2}{1+a^2}$$

alcanza su valor mínimo cuando es $a = 0$.

2. Sea r' la proyección ortogonal de r sobre π; sea $\bar{a}$ un vector unitario ortogonal a π, sea $\bar{u}'$ un vector unitario de la dirección de r' (nótese que $\bar{a} \perp \bar{u}'$), con lo que un vector $\bar{u}$ de la dirección de r lo será el $\bar{u} = \alpha\bar{a} + \bar{u}'$, para un cierto $\alpha \in \mathbb{R}$. Sea $\bar{v}$ un vector unitario ortogonal a $\bar{a}$ y a $\bar{u}'$. Hay que comprobar que $|\cos(\bar{u}, \bar{u}')|$ es el mayor de todos los $|\cos(\bar{u}, \bar{u}' + \lambda\bar{v})|$ para $\lambda \in \mathbb{R}$. Así es, ya que

$$|\cos(\bar{u}, \bar{u}' + \lambda\bar{v})| = \frac{1}{\sqrt{\alpha^2+1}\sqrt{1+\lambda^2}} \leqslant \frac{1}{\sqrt{\alpha^2+1}\sqrt{1}} = |\cos(\bar{u}, \bar{u}')|$$

DISTANCIA DE UN PUNTO A UN PLANO

[229]

Sean dados, en el espacio euclídeo E_3, un punto P y un plano π. Si P no pertenece a π, entonces existe un punto P_0 de π tal que el vector $\overrightarrow{PP_0}$ es ortogonal a π; se dice que P_0 es la *proyección ortogonal* de P sobre π. La distancia de P a P_0 es la menor de las distancias de P a los puntos de π y se la llama distancia del punto P al plano π.

Si, respecto de una referencia rectangular, $P(p_1, p_2, p_3)$ y π tiene por ecuación a $a_1x_1 + a_2x_2 + a_3x_3 + a_0 = 0$, entonces la distancia de P a π vale:

$$d(P, r) = \frac{|a_1p_1 + a_2p_2 + a_3p_3 + a_0|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}$$

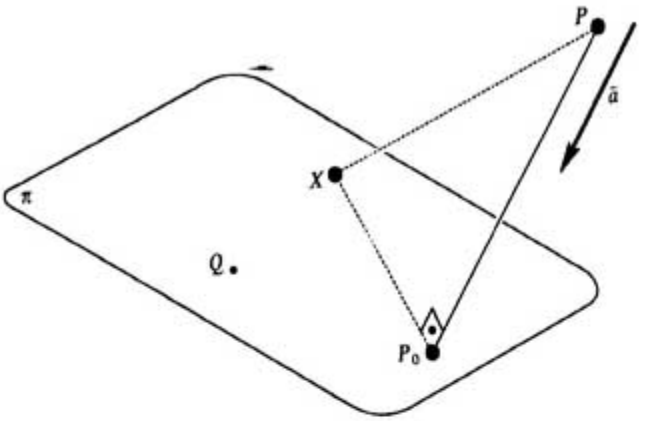


Figura 11.25.

DEMOSTRACION

Vamos a recurrir a un punto $Q \in \pi$ y a un vector $\bar{a}$ ortogonal a π. Hemos de comprobar que existe un $\lambda \in \mathbb{R}$ tal que para el punto $P_0 = P + \lambda\bar{a}$ se verifica que $\overrightarrow{QP_0} \cdot \bar{a} = 0$. Poniendo $\overrightarrow{QP_0} = \overrightarrow{PP_0} - \overrightarrow{PQ}$, el λ buscado es, pues, aquel que verifica a

$$(\lambda\bar{a} - \overrightarrow{PQ}) \cdot \bar{a} = 0, \quad \text{luego} \quad \lambda = \frac{\overrightarrow{PQ} \cdot \bar{a}}{\|\bar{a}\|^2}$$

La proyección ortogonal de P sobre π es el punto $P_0 = P + \lambda\bar{a}$ y la distancia $d(P, P_0)$ vale:

$$d(P, P_0) = \|\lambda\bar{a}\| = |\lambda|\,\|a\| = \frac{|\overrightarrow{PQ} \cdot \bar{a}|}{\|\bar{a}\|} \qquad [1]$$

Para cualquiera que sea el punto $X \neq P_0$ de π, como el triángulo PP_0X es rectángulo en P_0 y según el teorema de Pitágoras (véase [200]), se verifica que

$$\|\overrightarrow{PX}\|^2 = \|\overrightarrow{PP_0}\|^2 + \|\overrightarrow{P_0X}\|^2 > \|\overrightarrow{PP_0}\|^2, \quad \text{luego} \quad \|\overrightarrow{PX}\| > \|\overrightarrow{PP_0}\|$$

Por tanto, $d(P, P_0) = \|\overrightarrow{PP_0}\|$ es la menor de las distancias de P a los puntos de π.

Finalmente, recurriendo a coordenadas rectangulares, como podemos tomar $\bar{a}(a_1, a_2, a_3)$ y es $P(p_1, p_2, p_3)$ y teniendo en cuenta que, por ser $Q(q_1, q_2, q_3)$ un punto de π, se verifica que $a_1q_1 + a_2q_2 + a_3q_3 = -a_0$, se obtiene que:

$$d(P, r) = d(P, P_0) = \frac{|\overrightarrow{PQ} \cdot \bar{a}|}{\|\bar{a}\|} = \frac{|(q_1 - p_1)a_1 + (q_2 - p_2)a_2 + (q_3 - p_3)a_3|}{\sqrt{a_1^2 + a_2^2 + a_3^2}} =$$

$$= \frac{|a_1p_1 + a_2p_2 + a_3p_3 + a_0|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}$$

EJERCICIOS

1. Dados dos planos paralelos π y π', cuyas ecuaciones serán, pues, del tipo (en coordenadas rectangulares):

$$\pi: a_1x_1 + a_2x_2 + a_3x_3 + a_0 = 0 \quad \text{y} \quad \pi': a_1x_1 + a_2x_2 + a_3x_3 + a_0' = 0$$

pruébese que la distancia de un punto cualquiera de π a π' es constante. Esta distancia, que se llama distancia de π a π', vale:

$$d(\pi, \pi') = \frac{|a_0 - a_0'|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}$$

2. Sean r y π una recta y un plano paralelos. Compruébese que la distancia de un punto cualquiera de r a π es constante; a esta distancia se la llama distancia de r a π.

RESOLUCION

1. Sea $P(p_1, p_2, p_3)$ un punto cualquiera del plano π, esto es, tal que se verifica $a_1p_1 + a_2p_2 + a_3p_3 + a_0 = 0$. La distancia $d(P, \pi')$ vale, según se acaba de ver:

$$d(P, \pi') = \frac{|a_1p_1 + a_2p_2 + a_3p_3 + a_0'|}{\sqrt{a_1^2 + a_2^2 + a_3^2}} = \frac{|-a_0 + a_0'|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}$$

que es constante (no depende de $P \in \pi$) y tiene el valor que se dijo en el enunciado.

2. Sea τ el plano paralelo a π y que contiene a r. Si P recorre r, la distancia $d(P, \pi)$, como es la distancia de un punto de τ a π, según se acaba de decir es constante, no depende de P.

DISTANCIA DE UN PUNTO A UNA RECTA

[230]

Sean dados, en el espacio euclídeo E_3, un punto P y una recta r. Si r no pasa por P, entonces existe un punto P_0 de r tal que el vector $\overrightarrow{PP_0}$ es ortogonal a r; se dice que P_0 es la *proyección ortogonal* de P sobre r. La distancia de P a P_0 es la menor de las distancias de P a los puntos de r y se la llama distancia del punto P a la recta r.

Si r es la recta $Q + V(\bar{u})$, entonces la distancia de P a r vale:

$$d(P, r) = \frac{\|\overrightarrow{PQ} \wedge \bar{u}\|}{\|\bar{u}\|}$$

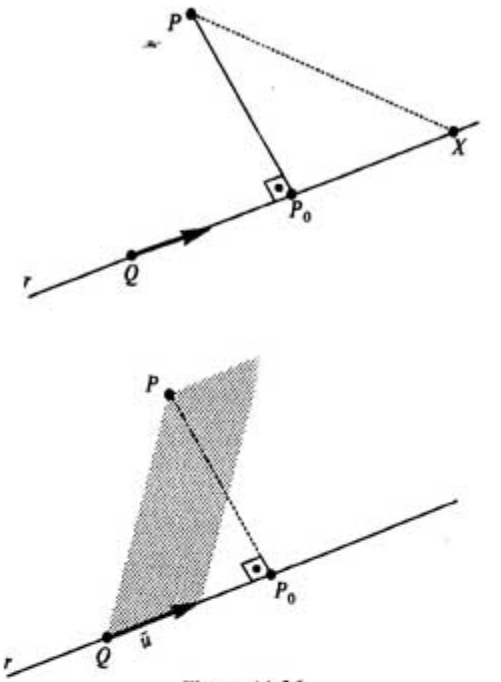


Figura 11.26.

DEMOSTRACION

Vamos a recurrir a un punto $Q \in r$ y a un vector $\bar{u} \neq \bar{o}$ de la dirección de r. Hemos de comprobar que existe un $\lambda \in \mathbb{R}$ tal que para el punto $P_0 = Q + \lambda\bar{u}$ se verifique $\overrightarrow{PP_0} \cdot \bar{u} = 0$. Poniendo $\overrightarrow{PP_0} = \overrightarrow{PQ} + \overrightarrow{QP_0}$, el λ buscado es, pues, aquel que verifica a:

$$(\overrightarrow{PQ} + \lambda\bar{u}) \cdot \bar{u} = 0, \quad \text{luego} \quad \lambda = -\frac{\overrightarrow{PQ} \cdot \bar{u}}{\|\bar{u}\|^2}$$

Por tanto, existe la proyección ortogonal, $P_0 = Q + \lambda\bar{u}$, de P sobre r.

Para cualquiera que sea el punto $X \neq P_0$ de r, como el triángulo PP_0X es rectángulo en P_0 y según el teorema de Pitágoras (véase [200]), se verifica que:

$$\|\overrightarrow{PX}\|^2 = \|\overrightarrow{PP_0}\|^2 + \|\overrightarrow{P_0X}\|^2 > \|\overrightarrow{PP_0}\|^2, \quad \text{luego} \quad \|\overrightarrow{PX}\| < \|\overrightarrow{PP_0}\|$$

Así, pues, $d(P, P_0) = \|\overrightarrow{PP_0}\|$ es la menor de las distancias de P a los puntos de r.

Finalmente, para obtener una expresión de la distancia $d(P, r)$, recurriendo al producto vectorial $\overrightarrow{PQ} \wedge \bar{u}$, se tiene:

$$\overrightarrow{PQ} \wedge \bar{u} = (\overrightarrow{PP_0} + \overrightarrow{P_0Q}) \wedge \bar{u} = \overrightarrow{PP_0} \wedge \bar{u}$$

(pues $\overrightarrow{P_0Q} \wedge \bar{u} = \bar{o}$, por ser $\overrightarrow{P_0Q}$ y $\bar{u}$ proporcionales). Por tanto, como $\overrightarrow{PP_0}$ y $\bar{u}$ son ortogonales, será:

$$\|\overrightarrow{PQ} \wedge \bar{u}\| = \|\overrightarrow{PP_0} \wedge \bar{u}\| = \|\overrightarrow{PP_0}\| \, \|\bar{u}\| \operatorname{sen} \frac{\pi}{2} = d(P, r)\|\bar{u}\|$$

de donde, al despejar $d(P, r)$, se obtiene la expresión buscada.

EJERCICIO

Sean r y r' dos rectas paralelas, de E_3. Compruébese que la distancia de un punto P cualquiera de r' a la recta r es constante (no depende de $P \in r'$); a esta distancia se la llama distancia entre r y r'.

RESOLUCION

Sea $\bar{u}$ un vector de la dirección de las rectas r y r'; sean Q y Q' puntos fijos de r y r'. Para cualquiera que sea $P \in r'$, se verifica que:

$$\overrightarrow{PQ} \wedge \bar{u} = (\overrightarrow{PQ'} + \overrightarrow{Q'Q}) \wedge \bar{u} = \bar{o} + \overrightarrow{Q'Q} \wedge \bar{u}$$

ya que $\overrightarrow{PQ'}$ y $\bar{u}$ son proporcionales. Por tanto,

$$d(P, r) = \frac{\|\overrightarrow{PQ} \wedge \bar{u}\|}{\|\bar{u}\|} = \frac{\|\overrightarrow{Q'Q} \wedge \bar{u}\|}{\|\bar{u}\|}$$

que no depende de P, como había que comprobar.

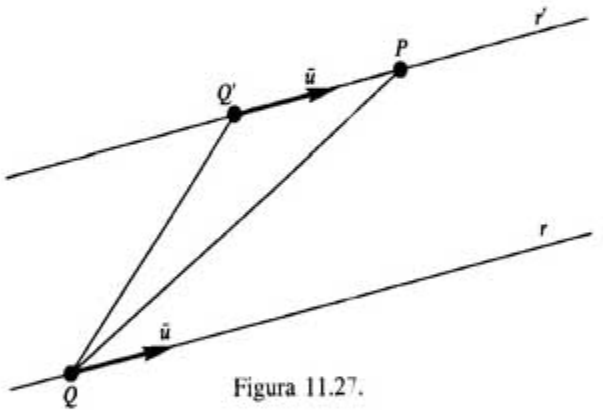


Figura 11.27.

DISTANCIA ENTRE DOS RECTAS QUE SE CRUZAN

[231]

Sean dadas dos rectas r y s, del espacio euclídeo E_3, que se cruzan. Se verifica que existe un punto $P_0 \in r$ y otro punto $Q_0 \in s$ tales que el vecto $\overrightarrow{P_0Q_0}$ es ortogonal a r y a s; a la recta P_0Q_0 e la llama *perpendicular común* a las rectas r y s. La distancia de P_0 a Q_0 es la menor de las distancias de un punto de r a un punto de s y se la llama distancia entre las rectas r y s.

Si r y s son las rectas $P + V(\bar{u})$ y $Q + V(\bar{v})$, entonces la distancia entre r y s vale:

$$d(r, s) = \frac{|\overrightarrow{PQ} \cdot (\bar{u} \wedge \bar{v})|}{\|\bar{u} \wedge \bar{v}\|}$$

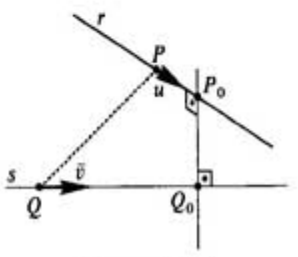


Figura 11.28.

DEMOSTRACION

Vamos a recurrir a dos puntos fijos $P \in r$ y $Q \in s$ y a dos vectores $\bar{u}$ y $\bar{v}$ de las direcciones de r y s; sean $\bar{u}_0$ y $\bar{v}_0$ vectores unitarios de las direcciones de $\bar{u}$ y $\bar{v}$, de manera que $\|\bar{u}_0\| = 1$, $\|\bar{v}_0\| = 1$ y $\bar{u}_0 \cdot \bar{v}_0 = \cos\theta$, siendo $\cos\theta \neq \pm 1$ (pues r y s no son paralelas). Hemos de comprobar que existen $\lambda, \mu \in \mathbb{R}$ tales que para los puntos $P_0 = P + \lambda\bar{u}_0$ y $Q_0 = Q + \mu\bar{v}_0$ se verifique $\overrightarrow{P_0Q_0} \cdot \bar{u}_0 = 0$ y $\overrightarrow{P_0Q_0} \cdot \bar{v}_0 = 0$. Poniendo $\overrightarrow{P_0Q_0} = \overrightarrow{PQ} - \lambda\bar{u}_0 - \mu\bar{v}_0$, los λ y μ buscados son, pues, aquellos que verifican a:

$$\left.\begin{aligned} 0 &= (\overrightarrow{PQ} - \lambda\bar{u}_0 + \mu\bar{v}_0) \cdot \bar{u}_0 = \overrightarrow{PQ} \cdot \bar{u}_0 - \lambda + \mu\cos\theta \\ 0 &= (\overrightarrow{PQ} - \lambda\bar{u}_0 + \mu\bar{v}_0) \cdot \bar{v}_0 = \overrightarrow{PQ} \cdot \bar{v}_0 - \lambda\cos\theta + \mu \end{aligned}\right\}$$

Este sistema de ecuaciones lineales (en las incógnitas λ y μ) tiene solución única, pues su determinante vale $1 - \cos^2\theta \neq 0$. Existen, pues, unos únicos puntos $P_0 \in r$ y $Q_0 \in s$ tales que $\overrightarrow{P_0Q_0}$ es ortogonal a r y a s.

Veamos ahora que, para cualesquiera $P' \in r$ y $Q' \in s$, se verifica que $d(P', Q') \geq d(P_0, Q_0)$. Para comprobarlo, recurriendo a que $P' = P_0 + \alpha\bar{u}$ y $Q' = Q_0 + \beta\bar{v}$ para ciertos $\alpha, \beta \in \mathbb{R}$, se tiene:

$$d(P', Q')^2 = \|\overrightarrow{P'Q'}\|^2 = \overrightarrow{P'Q'} \cdot \overrightarrow{P'Q'} =$$

$$= [\overrightarrow{P_0Q_0} + (-\alpha\bar{u} + \beta\bar{v})] \cdot [\overrightarrow{P_0Q_0} + (-\alpha\bar{u} + \beta\bar{v})] =$$

$$= \|\overrightarrow{P_0Q_0}\|^2 + \|-\alpha\bar{u} + \beta\bar{v}\|^2 \geq \|\overrightarrow{P_0Q_0}\|^2$$

(puesto que $\overrightarrow{P_0Q_0}$ es ortogonal a $-\alpha\bar{u} + \beta\bar{v}$); por tanto $\|\overrightarrow{P'Q'}\| \geq \|\overrightarrow{P_0Q_0}\|$, como había que comprobar.

Finalmente, para obtener la expresión que da la distancia $d(r, s) = \|\overrightarrow{P_0Q_0}\|$, vamos a recurrir al vector $\bar{u} \wedge \bar{v}$, que tiene la misma dirección que $\overrightarrow{P_0Q_0}$, lo que permite poner

$$|\overrightarrow{P_0Q_0} \cdot (\bar{u} \wedge \bar{v})| = \|\overrightarrow{P_0Q_0}\| \, \|\bar{u} \wedge \bar{v}\| = d(r, s)\|\bar{u} \wedge \bar{v}\| \qquad [1]$$

Ahora bien, como $\overrightarrow{P_0Q_0} = \overrightarrow{PQ} - \lambda\bar{u} + \mu\bar{v}$ para unos ciertos $\lambda, \mu \in \mathbb{R}$ y como $\bar{u} \wedge \bar{v}$ es ortogonal a $\bar{u}$ y a $\bar{v}$, el primer miembro de [1] se puede poner:

$$|\overrightarrow{P_0Q_0} \cdot (\bar{u} \wedge \bar{v})| = |(\overrightarrow{PQ} - \lambda\bar{u} + \mu\bar{v}) \cdot (\bar{u} \wedge \bar{v})| = |\overrightarrow{PQ} \cdot (\bar{u} \wedge \bar{v}) + \bar{o}|$$

llevando este último resultado a [1] y despejando $d(r, s)$, se obtiene la expresión buscada.

«PERPENDICULAR COMUN» A DOS RECTAS QUE SE CRUZAN

[232]

Dadas dos rectas que se cruzan, del espacio euclídeo E_3, existe una recta perpendicular a ambas y que se corta con ellas; esta recta se llama «perpendicular común» de aquellas.
Si las rectas dadas son $r = P + V(\vec{u})$ y $s = Q + V(\vec{v})$, entonces su perpendicular común es la intersección de los siguientes planos:

$$\begin{cases} X = P + \lambda_1\vec{u} + \mu_1(\vec{u} \wedge \vec{v}) \\ X = Q + \lambda_2\vec{v} + \mu_2(\vec{u} \wedge \vec{v}) \end{cases} \quad (\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R} \text{ parámetros})$$

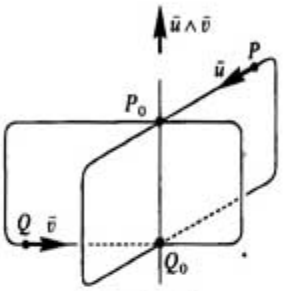


Figura 11.29.

DEMOSTRACION

La perpendicular común es la recta P_0Q_0 del apartado anterior (que existe según se vio allí), ya que $P_0 \in r$, $Q_0 \in s$ y $\overrightarrow{P_0Q_0}$ es ortogonal a r y a s. Es evidente que la recta P_0Q_0 es la intersección del plano que pasa por P, P_0 y Q_0 con el plano que pasa por Q, Q_0 y P_0. Como los vectores $\overrightarrow{PP_0}$, $\overrightarrow{QQ_0}$ y $\overrightarrow{P_0Q_0}$ tienen las direcciones de los vectores $\vec{u}$, $\vec{v}$ y $\vec{u} \wedge \vec{v}$, respectivamente, resulta que las direcciones de los dos planos anteriores son $V(\vec{u}, \vec{u} \wedge \vec{v})$ y $V(\vec{v}, \vec{u} \wedge \vec{v})$, luego dichos planos han resultado ser los del enunciado, esto es, los:

$$P + V(\vec{u}, \vec{u} \wedge \vec{v}) \quad \text{y} \quad Q + V(\vec{v}, \vec{u} \wedge \vec{v})$$

EJERCICIO

Hallar la distancia entre las siguientes rectas r y s (en coordenadas rectangulares):

$$r: \frac{x-1}{-1} = \frac{y-1}{3} = \frac{z+6}{2} \qquad s \begin{cases} x + y - 2z = 3 \\ x - 2y + 2z = 0 \end{cases}$$

RESOLUCION

La recta r pasa por $P(1, 1, -6)$ y tiene la dirección del vector $\vec{u}(-1, 3, 2)$; la recta s pasa por $Q(2, 1, 0)$ y tienen la dirección del vector $\vec{v}(2, 4, 3)$. Por tanto, la distancia $d(r, s)$ valdrá

$$d(r, s) = \frac{|\overrightarrow{PQ} \cdot (\vec{u} \wedge \vec{v})|}{\|\vec{u} \wedge \vec{v}\|} = \frac{|(1, 0, 6) \cdot [(-1, 3, 2) \wedge (2, 4, 3)]|}{\|(-1, 3, 2) \wedge (2, 4, 3)\|} =$$

$$= \frac{|1 + 0 - 60|}{\sqrt{1^2 + 7^2 + 10^2}} = \frac{59}{\sqrt{150}} \simeq 4{,}82$$

ECUACION NORMAL DE UN PLANO

[233]

Sea π un plano del espacio euclídeo E_3; en E_3 se ha adoptado una referencia cartesiana rectangular $(O; \bar{e}_1, \bar{e}_2, \bar{e}_3)$ y se supone que π no pasa por O. Llamando O_π a la proyección ortogonal de O sobre π y si $\overrightarrow{OO_\pi}(d \cos \theta_1, d \cos \theta_2, d \cos \theta_3)$, donde $d > 0$ es la distancia de O a π y $\theta_i = \text{áng}\,(\overrightarrow{OO_\pi}, \bar{e}_i)$, entonces π admite por ecuación a:

$$x_1 \cos \theta_1 + x_2 \cos \theta_2 + x_3 \cos \theta_3 = d$$

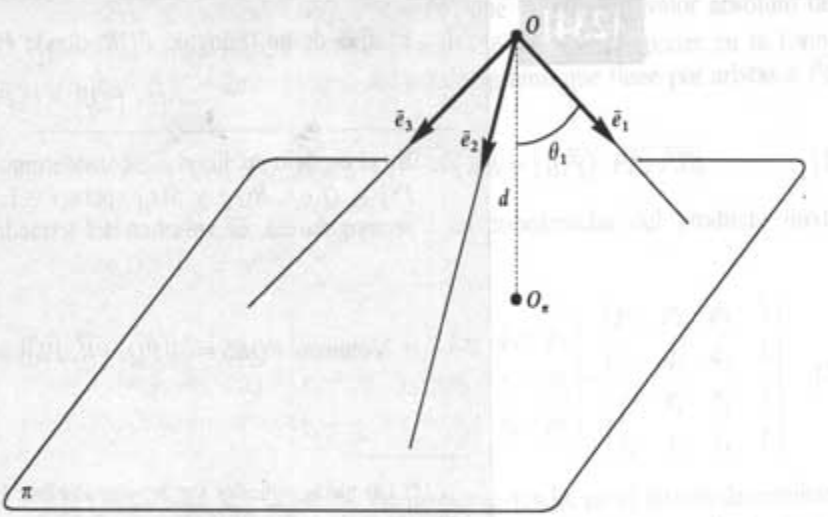


Figura 11.30.

COMPROBACION

Un punto X es de π si y sólo si $\overrightarrow{O_\pi X} \cdot \overrightarrow{OO_\pi} = 0$, esto es, si:

$$(x_1 - d \cos \theta_1) d \cos \theta_1 + (x_2 - d \cos \theta_2) d \cos \theta_2 + (x_3 - d \cos \theta_3) d \cos \theta_3 = 0$$

como $\cos^2 \theta_1 + \cos^2 \theta_2 + \cos^2 \theta_3 = 1$, ya que $(\cos \theta_2, \cos \theta_2, \cos \theta_3)$ es un vector unitario, la anterior ecuación resulta ser la del enunciado.

EJERCICIO

Sean r y s dos rectas que se cruzan. Obtener una referencia cartesiana rectangular $(O; \bar{e}_1, \bar{e}_2, \bar{e}_3)$ en la que las ecuaciones de r y s sean del tipo (llamamos x, y, z a las coordenadas)

$$r: \begin{cases} y = mx \\ z = a \end{cases} \quad \text{y} \quad s: \begin{cases} y = -mx \\ z = -a \end{cases}$$

RESOLUCION

El «eje de la z» es la perpendicular común a r y s (sean P y Q los puntos de corte de esta perpendicular con r y s); el origen es el punto medio de los P y Q; $\bar{e}_3$ tiene el sentido de $\overrightarrow{OP}$. Si r' y s' son las paralelas a r y s por O, las bisectrices de r' y s' son los «ejes de la x y de la y».

El número a es la mitad de la distancia $d(r, s)$; es $m = \operatorname{tg} \theta/2$, donde $\theta = \text{áng}\,(r, s)$.

AREA DE UN TRIANGULO; VOLUMEN DE UN TETRAEDRO

[234]

1. El área de un triángulo PQR, donde $P, Q, R \in E_3$ (euclídeo), vale:

$$\text{Area } \overset{\triangle}{PQR} = \tfrac{1}{2}\|\overrightarrow{PQ} \wedge \overrightarrow{PR}\|$$

2. En el espacio euclídeo E_3, consideremos cuatro puntos no coplanarios $P(p_i)$, $Q(q_i)$, $R(r_i)$ y $S(s_i)$ (para $i = 1, 2, 3$; estas coordenadas son rectangulares). El volumen del tetraedro $PQRS$ vale:

$$\text{Volumen } PQRS = \tfrac{1}{6}|[\overrightarrow{PQ}, \overrightarrow{PR}, \overrightarrow{PS}]| = \tfrac{1}{6}\left|\det \begin{bmatrix} p_1 & p_2 & p_3 & 1 \\ q_1 & q_2 & q_3 & 1 \\ r_1 & r_2 & r_3 & 1 \\ s_1 & s_2 & s_3 & 1 \end{bmatrix}\right| \quad (*)$$

(*) Las barras verticales son de valor absoluto. La expresión $[\overrightarrow{PQ}, \overrightarrow{PR}, \overrightarrow{PS}]$ denota el producto mixto de $\overrightarrow{PQ}$, $\overrightarrow{PR}$ y $\overrightarrow{PS}$.

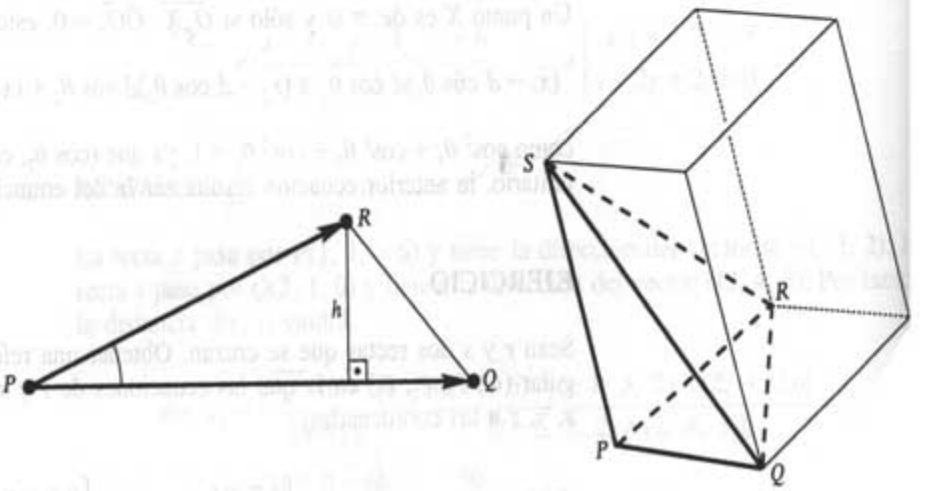


Figura 11.31.

COMPROBACION

1. La altura del triángulo correspondiente el vértice R se puede poner en la forma:

$$h = \|\overrightarrow{PR}\| \text{ sen } (\overrightarrow{PQ}, \overrightarrow{PR})$$

Recurriendo a la expresión que da la norma de un producto vectorial (véase [158], II), el área del triángulo, PQR, se puede poner:

$$\text{Area } \overset{\triangle}{PQR} = \tfrac{1}{2}\|\overrightarrow{PQ}\|h = \tfrac{1}{2}\|\overrightarrow{PQ}\|\,\|\overrightarrow{PR}\| \text{ sen } (\overrightarrow{PQ}, \overrightarrow{PR}) = \tfrac{1}{2}\|\overrightarrow{PQ} \wedge \overrightarrow{PR}\|$$

2. El volumen del tetraedro es $(1/3)bh$, donde b es el área del triángulo PQR y h es la altura correspondiente al vértice S. El volumen del paralelepípedo que tiene por aristas a $\overrightarrow{PQ}$, $\overrightarrow{PR}$ y $\overrightarrow{PS}$, que es igual al valor absoluto del producto mixto $[\overrightarrow{PQ}, \overrightarrow{PR}, \overrightarrow{PS}]$ (véase [156]), se puede poner en la forma $b'h$, donde $b' = 2b$ es el área del paralelogramo que tiene por aristas a $\overrightarrow{PQ}$ y $\overrightarrow{PR}$. Por tanto:

$$\text{Volumen } PQRS = \tfrac{1}{3}bh = \tfrac{1}{6}b'h = \tfrac{1}{6}|[\overrightarrow{PQ}, \overrightarrow{PR}, \overrightarrow{PS}]| \qquad [1]$$

Echando mano de la expresión en coordenadas del producto mixto (véase [154]), se tiene:

$$[\overrightarrow{PQ}, \overrightarrow{PR}, \overrightarrow{PS}] = \begin{vmatrix} q_1 - p_1 & q_2 - p_2 & q_3 - p_3 \\ r_1 - p_1 & r_2 - p_2 & r_3 - p_3 \\ s_1 - p_1 & s_2 - p_2 & s_3 - p_3 \end{vmatrix} = \begin{vmatrix} p_1 & p_2 & p_3 & 1 \\ q_1 & q_2 & q_3 & 1 \\ r_1 & r_2 & r_3 & 1 \\ s_1 & s_2 & s_3 & 1 \end{vmatrix} \qquad [2]$$

(la última igualdad se puede comprobar restando, en el último determinante, la primera fila a cada una de las demás y, luego, desarrollando por la última columna el determinante así obtenido). Llevando este resultado [2] a la expresión [1], se obtiene la fórmula que había que comprobar.

EJERCICIO

Sean r y s dos rectas que se cruzan. Sobre r se toman dos puntos P y Q; sobre s se toman dos puntos R y S. Hállese el volumen V del tetraedro cuyos vértices son P, Q, R y S, en función de

$$a = d(P, Q), \quad b = d(R, s), \quad \delta = d(r, s) \quad \text{y} \quad \alpha = \text{áng}\,(r, s)$$

RESOLUCION

Si $\bar{u}$ y $\bar{v}$ son vectores de la dirección de r y s, respectivamente, tales que $\|\bar{u}\| = a$ y $\|\bar{v}\| = b$, de acuerdo con [234], 2 y recurriendo a la expresión que da δ (véase [231]), se tiene:

$$V = \tfrac{1}{6}|\overrightarrow{PQ} \cdot (\overrightarrow{PR} \wedge \overrightarrow{PS})| = \tfrac{1}{6}|\bar{u} \cdot [\overrightarrow{PR} \wedge (\overrightarrow{PR} + \bar{v})]| = \tfrac{1}{6}|\bar{u} \cdot (\overrightarrow{PR} \wedge \bar{v})| =$$
$$= \tfrac{1}{6}|[\overrightarrow{PQ}, \bar{u}, \bar{v}]| = \tfrac{1}{6}\delta\|\bar{u} \wedge \bar{v}\| = \tfrac{1}{6}\delta ab \text{ sen } \alpha$$

LAS ESFERAS

[235]

(I) En el espacio euclídeo E_3, se llama *esfera*, de centro en el punto $C \in E_3$ y cuyo radio es $\rho > 0$, al lugar geométrico de los puntos de E_3 que distan ρ de C. Adoptando una referencia cartesiana rectangular en E_3 y si $C(a, b, c)$, dicha circunferencia está formada por los puntos $X(x, y, z)$ tales que:

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = \rho^2$$

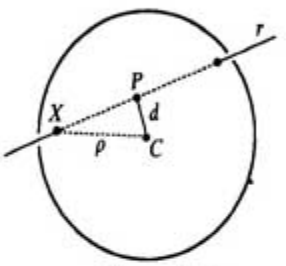


Figura 11.32.

(II) La esfera anterior y una recta r se cortan en dos, uno o ningún punto, según que la distancia d de C a r sea $d < \rho$, $d = \rho$ o $d > \rho$ y se dice que r es secante, *tangente* o exterior a la esfera, respectivamente.

Por un punto $P(x_0, y_0, z_0)$ de la esfera pasan infinitas tangentes a ésta; todas ellas están situadas en el plano perpendicular a CP por P, el cual tiene por ecuación a

$$(x_0 - a)(x - x_0) + (y_0 - b)(y - y_0) + (z_0 - c)(z - z_0) = 0$$

A éste, se le llama *plano tangente* a la espera en el punto P de ella.

DEMOSTRACION

(I) Un punto $X(x, y, z)$ será de la esfera si y sólo si $\|\overrightarrow{CX}\| = \rho$, esto es, si $\|\overrightarrow{CX}\|^2 = \rho^2$: como $\overrightarrow{CX}(x-a,\ y-b,\ z-c)$, la anterior relación puede ponerse en la forma:

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = \rho^2$$

(II) Llamando P a la proyección ortogonal de C sobre r y si $\bar{u}$ es un vector unitario de la dirección de r, los puntos de esta recta r serán los $X = P + \lambda\bar{u}$ para $\lambda \in \mathbb{R}$. Un punto $X \in r$ será de la esfera cuando $\|\overrightarrow{CX}\| = \rho$; como el triángulo CPX es rectángulo en P y según el teorema de Pitágoras (véase [200]), la última relación es equivalente a:

$$\rho^2 = \|\overrightarrow{CX}\|^2 = \|\overrightarrow{CP}\|^2 + \|\overrightarrow{PX}\|^2 = d^2 + \lambda^2 \quad \text{o} \quad \lambda^2 = \rho^2 - d^2$$

Por tanto, según que sea $\rho > d$, $\rho = d$ o $\rho < d$, habrá dos, una o ninguna solución para λ, esto es, dos, un o ningún punto $(X = P + \lambda\bar{u})$ de intersección. Nótese que el caso de tangencia (la intersección es un punto único) se presenta si y sólo si la distancia de C a r es ρ y, en tal caso, el punto P de tangencia (intersección) es la proyección ortogonal de C sobre r, de manera que $\overrightarrow{CP}$ es ortogonal a r.

Dado un punto $P(x_0, y_0, z_0)$ de la esfera, las tangentes a la esfera en P

son, pues, todas las rectas que pasan por P y son ortogonales al vector $\overrightarrow{CP}(x_0 - a, y_0 - b, z_0 - c)$, esto es, que pasan por P y están en el plano de ecuación

$$(x_0 - a)(x - x_0) + (y_0 - b)(y - y_0) + (z_0 - c)(z - z_0) = 0$$

EJERCICIO

Sean A y B dos puntos, de E_3, diametralmente opuestos de una esfera (tales que su punto medio es el centro). Pruébese que un punto $P \in E_3$ es de la referida esfera si y sólo si $\overrightarrow{AP}$ y $\overrightarrow{BP}$ son ortogonales.

RESOLUCION

Llamemos C y ρ al centro y al radio de la esfera, de manera que $A = C + \rho\bar{u}$ y $B = C - \rho\bar{u}$ para un cierto vector unitario $\bar{u} \in \mathbb{R}^3$. La distancia $d(C, P)$ se puede poner:

$$\begin{aligned} d(C, P)^2 = \|\overrightarrow{CP}\|^2 = \overrightarrow{CP} \cdot \overrightarrow{CP} = (\overrightarrow{CA} + \overrightarrow{AP}) \cdot (\overrightarrow{CB} + \overrightarrow{BP}) = \\ = (\overrightarrow{AP} + \rho\bar{u}) \cdot (\overrightarrow{BP} - \rho\bar{u}) = \overrightarrow{AP} \cdot \overrightarrow{BP} + \rho(\overrightarrow{BP} - \overrightarrow{AP}) \cdot \bar{u} - \rho^2 = \\ = \overrightarrow{AP} \cdot \overrightarrow{BP} + \rho\overrightarrow{BA} \cdot \bar{u} - \rho^2 = \overrightarrow{AP} \cdot \overrightarrow{BP} + 2\rho^2\bar{u} \cdot \bar{u} - \rho^2 = \\ = \overrightarrow{AP} \cdot \overrightarrow{BP} + \rho^2 \end{aligned}$$

Por tanto, la distancia $d(C, P)$ valdrá ρ si y sólo si $\overrightarrow{AP} \cdot \overrightarrow{BP} = 0$, es decir, P pertenecerá a la esfera si y sólo si $\overrightarrow{AP}$ y $\overrightarrow{BP}$ son ortogonales.

INTERSECCION DE UNA ESFERA Y UN PLANO

[236]

En el espacio euclídeo E_3, se considera un plano π y la esfera de centro en C y de radio $\rho > 0$. Llamando P a la proyección ortogonal de C sobre π y si es d la distancia de C a π, entonces la intersección del plano π con la esfera es:

- Si $\rho > d$, la circunferencia de π que tiene centro en P y radio $\sqrt{\rho^2 - d^2}$.
- Si $\rho = d$, el punto P (el *plano tangente* a la esfera en P es π).
- Si $\rho < d$, el vacío (todos los puntos de π son exteriores a la esfera).

DEMOSTRACION

Llamemos $\bar{u}$ al vector unitario de la dirección y sentido de $\overrightarrow{CP}$, con lo que $\overrightarrow{CP} = d\bar{u}$. Los puntos de π son los $X = P + \lambda\bar{v}$, para $\lambda \in \mathbb{R}$ y donde $\bar{v} \in \mathbb{R}^3$ es un vector ortogonal a $\bar{u}$, que podemos suponer unitario. Los puntos de la intersección son los $X \in E_3$ que, siendo de la forma antes dicha, son tales que

$\|\overrightarrow{CX}\| = \rho$. Los puntos de esta intersección, son, pues, los que se obtienen para los valores de λ que verifican a:

$$\rho^2 = \|\overrightarrow{CX}\|^2 = \|\overrightarrow{CP} + \overrightarrow{PX}\|^2 = \|d\vec{u} + \lambda\vec{v}\|^2 = d^2 + \lambda^2 \quad \text{o sea}$$
$$\lambda^2 = \rho^2 - d^2$$

Como $|\lambda|$ es la distancia $\|\overrightarrow{PX}\|$, los puntos X de la intersección son, pues:

- Si $\rho > d$, los puntos X de π tales que $\|\overrightarrow{PX}\| = \sqrt{\rho^2 - d^2}$, que son los de la circunferencia del enunciado.
- Si $\rho = d$, con lo que $\lambda = 0$, solamente se obtiene el punto $X = P$.
- Si $\rho < d$, no hay puntos de intersección (todos los del plano distan de C más de ρ, esto es, son exteriores a la esfera).

EJERCICIO

En el espacio euclídeo E_3 y utilizando coordenadas cartesianas rectangulares x, y, z, se considera la esfera de centro $C(a, b, c)$ y radio $\rho > 0$. Hallar la relación a la que han de satisfacer los números α, β, γ, $\delta \in \mathbb{R}$ para que el siguiente plano π sea tangente a la esfera:

$$\pi: \alpha x + \beta y + \gamma z + \delta = 0$$

RESOLUCION

El plano π es tangente a la esfera si y sólo si $d(C, \pi) = \rho$, esto es:

$$\frac{|\alpha a + \beta b + \gamma c + \delta|}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}} = \rho$$

o lo que es igual:

$$(\alpha a + \beta b + \gamma c + \delta)^2 = (\alpha^2 + \beta^2 + \gamma^2)\rho^2$$

EJERCICIO

Hallar el centro y el radio de la esfera inscrita en el tetraedro cuyos vértices son los puntos $A(3, 0, 0)$, $B(0, 3, 0)$, $C(0, 0, 3)$ y $D(-1, -1, -1)$.

RESOLUCION

Las caras del tetraedro, es decir, los planos $\pi_1 = ABC$, $\pi_2 = ABD$, $\pi_3 = ACD$ y $\pi_4 = BCD$, tienen las siguientes ecuaciones:

$$\pi_1: x + y + z = 3 \qquad \pi_3: x - 5y + z = 3$$
$$\pi_2: x + y - 5z = 3 \qquad \pi_4: -5x + y + z = 3$$

El centro de la esfera, $P(\alpha, \beta, \gamma)$ ha de equidistar de los cuatro planos, por lo que debe ser:

$$\frac{\alpha + \beta + \gamma - 3}{\sqrt{3}} = \frac{\alpha + \beta + 5\gamma - 3}{3\sqrt{3}} = \frac{\alpha - 5\beta + \gamma - 3}{3\sqrt{3}} = \frac{-5\alpha + \beta + \gamma - 3}{3\sqrt{3}}$$

(nótese que P es interior al tetraedro y que el origen de coordenadas también lo es, por lo que los signos que se han tomado para las anteriores distancias son correctos). De estas ecuaciones se obtiene que $\alpha = \beta = \gamma = 1/2$, o sea, $P(1/2, 1/2, 1/2)$. Finalmente, el radio de la esfera es la distancia de P a π, luego vale $\rho = \sqrt{3}/2$.

δ ESPACIO AFIN AMPLIADO $\bar{E}_n$ (en los casos $n = 2$ y $n = 3$)

En los espacios afines E_2 y E_3, al estudiar las intersecciones de las rectas (subespacios unidimensionales) y las de los planos (subespacios bidimensionales), nos hemos encontrado con una situación nueva, que no se presentaba al hallar las intersecciones de subespacios vectoriales; nos referimos al paralelismo. A este respecto, recordemos que dos rectas distintas de E_2 se cortan, en un punto de E_2, salvo que sean paralelas y que dos planos distintos de E_3 se cortan, según una recta de E_3, salvo si son paralelos. Para evitar estas excepciones, que en muchos momentos son francamente perturbadoras, vamos a ampliar los espacios E_2 y E_3, añadiéndoles nuevos puntos, que llamaremos «puntos impropios» o «puntos del infinito». Las cosas se harán de manera que a cada recta se le añada un nuevo punto, obteniéndose así una «recta ampliada»; tales puntos coinciden si las rectas son paralelas y son distintos en caso contrario. Con ello, dos rectas paralelas pasan, ahora, a tener un punto en común (un punto impropio o del infinito). Las cosas funcionan de modo que a un plano se le añade un punto impropio por cada una de sus direcciones; se dice que los puntos impropios de un plano forman una recta impropia o recta del infinito. Dos planos paralelos pasan, así, a tener en común una recta, que es su recta impropia.

11.8. COORDENADAS HOMOGENEAS

Los puntos impropios o del infinito deben introducirse de modo que admitan tratamiento analítico; se quiere que puedan determinarse mediante algún tipo de coordenadas. Es obvio que para nada vale decir que los puntos impropios son aquellos que tienen sus coordenadas valiendo infinito; además de que «infinito» (∞) no es un número, ello no permitiría, ni siquiera, distinguir a unos puntos impropios de otros. Así, pues, hay que idear un nuevo modo de asignar coordenadas con el que todos los puntos, propios e impropios, queden determinados por un adecuado sistema de números.

Como se desea introducir un punto del infinito por cada dirección y (en E_3, por ejemplo, y usando coordenadas cartesianas $(O-x_2x_2x_3)$ sabemos que dos puntos alineados con el origen, $X(x_1, x_2, x_3)$ y $X'(hx_1, hx_2, hx_3)$, determinan la misma dirección (pues $\overrightarrow{OX}$ y $\overrightarrow{OX'}$ son proporcionales), se puede decir que cualquiera de las ternas (hx_1, hx_2, hx_3), al variar h $(h \neq 0)$, representa a un mismo punto impropio. Para armonizar esto con lo que acontece con los puntos propios, de manera que dos sistemas de coordenadas que sean proporcionales representen, en todos los casos, a un mismo punto, se añade una nueva coordenada (llamada homogénea o de homogeneización) del siguiente modo:

COORDENADAS HOMOGENEAS DE UN PUNTO

[237]

En el espacio afín E_n (para $n = 2$ o $n = 3$)(*), se considera una referencia cartesiana $(O; \bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$. Si las coordenadas cartesianas de un punto $X \in E_n$ son $(x_1, x_2, ..., x_n)$, se dice que $(\rho, \rho x_1, \rho x_2, ..., \rho x_n)$ es, para cualquiera que sea $\rho \neq 0$, un sistema (sucesión) de *coordenadas homogéneas* de X; esto es, una sucesión $(x_0^*, x_1^*, x_2^*, ..., x_n^*)$, de $n + 1$ números reales, es un sistema de coordenadas homogéneas del punto X si

$$x_0^* \neq 0, \frac{x_1^*}{x_0^*} = x_1, \frac{x_2^*}{x_0^*} = x_2, ..., \frac{x_n^*}{x_0^*} = x_n$$

Con este criterio, un vector cualquiera $\bar{x} = (x_0^*, x_1^*, x_2^*, ..., x_n^*) \in \mathbb{R}^{n+1}$, con $x_0^* \neq 0$, define a un y sólo un punto $X \in E_n$. Cuando así ocurra, se pondrá $X(\rho x_i^*)$ o $X(\rho \bar{x})$, donde $\rho \neq 0$ representa a un número real arbitrario.

(*) Cuanto aquí se dice es de aplicación, no obstante, al caso de dimensión finita $n \in \mathbb{N}$ cualquiera.

OBSERVACIONES

1. Cuando se hable de las coordenadas cartesianas (x_i) de un punto y se quiera señalar que dichas coordenadas no son las coordenadas homogéneas (x_i^*), a aquéllas se las suele llamar coordenadas absolutas. Cuando sólo se manejan coordenadas homogéneas, se simplifica la notación y a ellas se las denota por $(x_0, x_1, x_2, ..., x_n)$, omitiendo el * que las diferencia de las coordenadas absolutas.
2. Cuando las coordenadas cartesianas (absolutas) se denotan por (x, y), en el plano E_2, o por (x, y, z), en el espacio E_3, a las coordenadas homogéneas se las suele llamar (T, X, Y) o (T, X, Y, Z), respectivamente, y también (t, x, y) o (t, x, y, z), cuando no quepa confusión con las coordenadas absolutas.
3. Un punto $X \in E_3$, fijo, tiene infinitos sistemas de coordenadas homogéneas (en una misma referencia cartesiana); todos ellos son proporcionales entre sí. Por ello, para señalar que $\bar{x} = (x_i^*) \in \mathbb{R}^{n+1}$ es un sistema de coordenadas homogéneas de X, se puso $X(\rho x_i^*)$ o $X(\rho \bar{x})$, en lugar de $X(x_i^*)$ o $X(\bar{x})$, donde ρ representa a un número real cualquiera no nulo.
4. Si las coordenadas cartesianas (absolutas) de un punto $X \in E_n$ son $(x_1, x_2, ..., x_n)$, entonces uno de sus sistemas de coordenadas homogéneas es el $(1, x_1, x_2, ..., x_n)$. Así, en E_3, el origen de coordenadas, y los puntos unidad de cada uno de los ejes, que tienen por coordenadas absolutas a $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$, admiten las siguientes sucesiones de coordenadas homogéneas $(1, 0, 0, 0)$, $(1, 1, 0, 0)$, $(1, 0, 1, 0)$ y $(1, 0, 0, 1)$, respectivamente; nótese que también son sucesiones de coordenadas homogéneas de los referidos puntos las siguientes $(4, 0, 0, 0)$, $(3, 3, 0, 0)$, $(-1, 0, -1, 0)$ y $(8, 0, 0, 8)$, respectivamente.

ECUACIONES GENERALES DE RECTAS Y PLANOS, EN HOMOGENEAS

[238]

En el plano afín E_2, en el que se ha adoptado una cierta referencia cartesiana (coordenadas x_1, x_2), consideramos la recta cuya ecuación es $a_0 + a_1x_1 + a_2x_2 = 0$, donde $(a_1, a_2) \neq (0, 0)$. Al expresar esta recta en coordenadas homogéneas, ella admite por ecuación a $a_0x_0^* + a_1x_1^* + a_2x_2^* = 0$.

En el espacio afín E_3, en el que se ha adoptado una cierta referencia cartesiana (coordenadas x_1, x_2, x_3), consideramos el plano cuya ecuación es $a_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0$, donde $(a_1, a_2, a_3) \neq (0, 0, 0)$. Al expresar este plano en coordenadas homogéneas, él admite por ecuación a $a_0x_0^* + a_1x_1^* + a_2x_2^* + \ + a_3x_3^* = 0$.

COMPROBACION

Consideremos el caso del plano de E_3; el de la recta de E_2 es aún más simple. Como $x_i = x_i^*/x_0^*$ (no se olvide que $x_0^* \neq 0$) para $i = 1, 2, 3$, es evidente que

$$a_0 + \sum_{i=1}^{3} a_ix_i = 0 \Leftrightarrow a_0 + \sum_{i=1}^{3} a_i \frac{x_i^*}{x_0^*} = 0 \Leftrightarrow a_0x_0^* + \sum_{i=1}^{3} a_ix_i^* = 0 \Leftrightarrow \sum_{i=0}^{3} a_ix_i^* = 0$$

OBSERVACIONES

1. Nótese que al utilizar coordenadas homogéneas en lugar de las absolutas, el único cambio que se aprecia en la ecuación es que la coordenada x_0^* aparece como factor en lo que era término independiente, que pasa así a ser de primer grado, como los otros sumandos. Es decir, las coordenadas homogéneas vienen a homogeneizar las ecuaciones de las rectas y los planos.
2. En las ecuaciones de las rectas y los planos en coordenadas homogéneas, la coordenada x_0^* aparece desempeñando un papel similar al de las otras coordenadas. Esto no es del todo así, pues en la condición $(a_1, a_2, a_3) \neq \neq (0, 0, 0)$ no interviene el coeficiente a_0.
   Imaginemos que, en la anterior ecuación homogénea (de la recta, si $n = 2$; del plano, si $n = 3$), se hiciesen nulos todos los coeficientes a_i excepto el a_0, nos encontraríamos con la relación $a_0x_0^* = 0$, o su equivalente $x_0^* = 0$, y esta relación no es satisfecha por ningún punto (de los que conocemos hasta el momento), ya que para todos ellos es $x_0^* \neq 0$.
3. Cuando, en el próximo capítulo, se consideren espacios afines E_n de dimensión finita $n \in \mathbb{N}$ cualquiera, podrá verse fácilmente que la ecuación $a_0x_0^* + a_1x_1^* + a_2x_2^* + \cdots + a_nx_n^* = 0$ (en coordenadas homogéneas x_0^*, x_1^*, x_2^*, ..., x_n^*) representa lo que se llama un *hiperplano* o variedad (subespacio) afín lineal de dimensión $n - 1$, siempre que $(a_1, a_2, ..., a_n) \neq (0, 0, ..., 0)$.

ECUACIONES PARAMETRICAS DE RECTAS Y PLANOS, EN HOMOGENEAS

[239]

En el espacio afín E_n (para $n = 2$ o $n = 3$)[1], expresando los puntos $X \in E_n$ a través de sus coordenadas homogéneas $(x_0^*, x_1^*, x_2^*, ..., x_n^*)$, respecto de una cierta referencia cartesiana, las rectas y los planos admiten las siguientes ecuaciones paramétricas:

- La *recta* que pasa por los puntos distintos $P(\rho p_i^*)$ y $Q(\rho q_i^*)$ es el conjunto $r(P, Q)$ de aquellos puntos $X(\rho x_i^*)$ tales que $\rho x_i^* = \alpha p_i^* + \beta q_i^*$ $(i = 0, 1, 2, ..., n)$ al variar $\alpha, \beta \in \mathbb{R}$, siendo $(\alpha, \beta) \neq (0, 0)$[2].
- El *plano* que pasa por los puntos no alineados $P(\rho p_i^*)$, $Q(\rho q_i^*)$ y $R(\rho r_i^*)$ es el conjunto $\pi(P, Q, R)$ que forman los puntos $X(\rho x_i^*)$ tales que $\rho x_i^* = \alpha p_i^* + \beta q_i^* + \gamma r_i^*$ $(i = 0, 1, 2, ..., n)$ al variar $\alpha, \beta, \gamma \in \mathbb{R}$, siendo $(\alpha, \beta, \gamma) \neq (0, 0, 0)$[2].

$$r(P, Q): \begin{cases} \rho x_0^* = \alpha p_0^* + \beta q_0^* \\ \rho x_1^* = \alpha p_1^* + \beta q_1^* \\ \rho x_2^* = \alpha p_2^* + \beta q_2^* \\ \cdots\cdots\cdots\cdots\cdots \\ \rho x_n^* = \alpha p_n^* + \beta q_n^* \end{cases} \quad \begin{matrix} (\alpha, \beta) \in \mathbb{R}^2 \\ (\alpha, \beta) \neq O \end{matrix}$$

$$\pi(P, Q, R) \begin{cases} \rho x_0^* = \alpha p_0^* + \beta q_0^* + \gamma r_0^* \\ \rho x_1^* = \alpha p_1^* + \beta q_1^* + \gamma r_1^* \\ \rho x_2^* = \alpha p_2^* + \beta q_2^* + \gamma r_2^* \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \rho x_n^* = \alpha p_n^* + \beta q_n^* + \gamma r_n^* \end{cases} \quad \begin{matrix} (\alpha, \beta, \gamma) \in \mathbb{R}^3 \\ (\alpha, \beta, \gamma) \neq O \end{matrix}$$

(1) Cuanto aquí se dice es válido, no obstante, para dimensión finita $n \in \mathbb{N}$ cualquiera.
(2) Se excluyen también los valores de (α, β) o de (α, β, γ) que conduzcan a $x_0^* = 0$.

DEMOSTRACION

Vamos a considerar solamente el caso del plano (plano de E_n, si $n > 2$); para la recta, el razonamiento es más simple: bastará con prescindir del punto R y del parámetro γ. Las relaciones $\rho x_i^* = \alpha p_i^* + \beta q_i^* + \gamma r_i^*$ (para $i = 0, 1, 2, ..., n$) se verifican para algunos valores de los parámetros α, β y γ (no todos nulos) si y sólo si

$$\operatorname{rang} \begin{bmatrix} x_0^* & x_1^* & x_2^* & \cdots & x_n^* \\ p_0^* & p_1^* & p_2^* & \cdots & p_n^* \\ q_0^* & q_1^* & q_2^* & \cdots & q_n^* \\ r_0^* & r_1^* & r_2^* & \cdots & r_n^* \end{bmatrix} = 3$$

Recurriendo a las coordenadas cartesianas (absolutas) x_i, p_i, q_i y r_i (para $i = 1, 2, ..., n$; recuérdese que $x_i = x_i^*/x_0^*$) la anterior relación equivale a

$$3 = \text{rang}\begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ 1 & p_1 & p_2 & \cdots & p_n \\ 1 & q_1 & q_2 & \cdots & q_n \\ 1 & r_1 & r_2 & \cdots & r_n \end{bmatrix} = \text{rang}\begin{bmatrix} 1 & p_1 & p_2 & \cdots & p_n \\ 0 & x_1 - p_1 & x_2 - p_2 & \cdots & x_n - p_n \\ 0 & q_1 - p_1 & q_2 - p_2 & \cdots & q_n - p_n \\ 0 & r_1 - p_1 & r_2 - p_2 & \cdots & r_n - p_n \end{bmatrix}$$

(la última matriz resulta de restar, en la primera, la segunda fila a todas las demás y, luego, permutar las dos primeras filas); ambas matrices tienen, pues, igual rango. En la última matriz, la primera fila es independiente de las otras tres, a causa del 1 que tiene en la primera columna, en la que los demás elementos son nulos. Por tanto, al prescindir de dicha fila primera, la anterior condición equivale a

$$\text{rang}\begin{bmatrix} x_1 - p_1 & x_2 - p_2 & \cdots & x_n - p_n \\ q_1 - p_1 & q_2 - p_2 & \cdots & q_n - p_n \\ r_1 - p_1 & r_2 - p_2 & \cdots & r_n - p_n \end{bmatrix} = 2$$

lo que significa que el vector $\overrightarrow{PX}(x_i - p_i)$ es combinación lineal de los $\overrightarrow{PQ}(q_i - p_i)$ y $\overrightarrow{PR}(r_i - p_i)$ (estos dos últimos son independientes, pues P, Q y R no están alineados), es decir, los puntos X que verifican esta relación son los del plano que pasa por P, Q y R como había que comprobar.

EJEMPLO

La recta de E_3 que pasa por los puntos $P(2, 1, 0, 4)$ y $Q(-1, 2, 1, 3)$ (en coordenadas homogéneas) tiene ecuaciones paramétricas:

$$x_0^* = 2\alpha - \beta \quad , \quad x_1^* = \alpha + 2\beta \quad , \quad x_2^* = \beta \quad , \quad x_3^* = 4\alpha + 3\beta \qquad [1]$$

Si eliminamos los parámetros α y β. obtenemos la recta como intersección de planos:

$$\begin{cases} x_0^* - 2x_1^* + 5x_2^* = 0 \\ 4x_1^* - 5x_2^* - x_3^* = 0 \end{cases}$$

Según sabemos, al dar valores a los parámetros α y β, las anteriores ecuaciones [1] nos proporcionan las coordenadas de los puntos de la recta. Nótese que si se toma $\alpha = \alpha_0$, $\beta = \beta_0$ y, luego, se toma $\alpha = \rho\alpha_0$, $\beta = \rho\beta_0$ (donde $\rho \neq 0$ es un escalar cualquiera), se obtiene el mismo punto, en uno y otro caso. También conviene observar que si se toma $\beta = 2\alpha$, se obtiene $x_0^* = 0$, que no es coordenada homogénea de ningún punto (de los conocidos hasta ahora); por tanto, los puntos de la recta se obtienen al tomar α y β de modo que $(\alpha, \beta) \neq (0, 0)$ y que $2\alpha - \beta \neq 0$.

CAMBIOS DE LA REFERENCIA EN HOMOGENEAS

[240]

Supongamos que, en el espacio afín E_n ($n = 2$ *o* $n = 3$)[1], se dispone de dos referencias cartesianas y que, según se verá (en [255]), las coordenadas (absolutas) (x_i) y (x_i') de un mismo punto $X \in E_n$, en una y otra referencia, guardan la siguiente relación[2]:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \cdots & \cdots & \cdots \\ q_{n1} & \cdots & q_{nn} \end{bmatrix} \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} + \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Abreviadamente:

$$X = QX' + C \qquad [1]$$

Entonces las coordenadas homogéneas (ρx_i^*) y $(\rho x_i^{*'})$ de X, en una y otra referencia, están relacionadas (para algún $\rho \neq 0$) mediante:

$$\rho \begin{bmatrix} x_0^* \\ x_1^* \\ \vdots \\ x_n^* \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline c_1 & q_{11} & \cdots & q_{1n} \\ \vdots & \cdots & \cdots & \cdots \\ c_n & q_{n1} & \cdots & q_{nn} \end{array}\right] \begin{bmatrix} x_0^{*'} \\ x_1^{*'} \\ \vdots \\ x_n^{*'} \end{bmatrix}$$

Abreviadamente:

$$\rho X^* = Q^* X^{*'} \qquad [2]$$

---

(1) Cuanto aquí se dice es válido en general, para dimensión finita $n \in \mathbb{N}$ cualquiera.

(2) Si las referencias son $(O; \bar{e}_1, \ldots, \bar{e}_n)$ y $(C; \bar{u}_1, \ldots, \bar{u}_n)$, se llaman (c_i) a las coordenadas de C en la primera referencia y q_{ij} a los escalares que permiten poner $\bar{u}_j = \sum_{i=1}^{n} q_{ij} \bar{e}_i$ $(i = 1, \ldots, n)$.

COMPROBACION

Para pasar de [1] a [2], téngase en cuenta que la relación [1] se puede expresar en la forma equivalente:

$$\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline c_1 & q_{11} & \cdots & q_{1n} \\ \vdots & \cdots & \cdots & \cdots \\ c_n & q_{n1} & \cdots & q_{nn} \end{array}\right] \begin{bmatrix} 1 \\ x_1' \\ \vdots \\ x_n' \end{bmatrix}$$

Si en esta última igualdad se pone $x_i = x_i^*/x_0^*$ y $x_i' = x_i^{*'}/x_0^{*'}$, luego se sacan $1 : x_0^*$ y $1 : x_0^{*'}$ como factores comunes de las matrices columna y se llama $\rho = x_0^{*'}/x_0^*$, se obtiene la relación [2].

11.9. ESPACIO AFIN AMPLIADO (O PROYECTIVIZADO)

Para llegar al concepto de espacio afín ampliado, vamos a proceder de manera que las definiciones vayan apareciendo de la mano de su interpretación intuitiva. Así, al imaginar que un punto variable se aleja, tendiendo al infinito, siguiendo una determinada dirección, se supondrá que ello conduce a una posición límite, que se llamará punto impropio y dependerá de la dirección según la que se aleje, y nada más que de ella. Las cosas deben, pues, funcionar de manera que todas las rectas que tienen la misma dirección, paralelas entre sí, terminen teniendo el mismo punto en el infinito.

PUNTOS IMPROPIOS O DEL INFINITO

[241]

En un espacio afín E_n ($n = 2$ o $n = 3$)(*), el paralelismo entre sus rectas es una relación de equivalencia (véase Apéndice A.4); las clases de esta equivalencia se pueden identificar con las direcciones de E_n, esto es, con los subespacios unidimensionales de $\mathbb{R}^n$. A cada una de estas clases se la llama *punto impropio o del infinito* de la dirección de todas sus rectas; si esta dirección es $V(\bar{u}) = \{\lambda\bar{u}/\lambda \in \mathbb{R}\}$ (donde $\bar{u} \in \mathbb{R}^n$ es no nulo cualquiera), al punto del infinito se le denotará también por $V(\bar{u})$ o poniendo $\bar{u}_\infty$.

(*) Cuanto aquí se dice es válido en general, para dimensión finita $n \in \mathbb{N}$ cualquiera.

COMPROBACION

La relación // (paralelismo), entre las rectas de E_n, se define por:

$$r//s \quad \Leftrightarrow \quad [(\text{dirección de } r) = (\text{dirección de } s)]$$

Es obvio que // es reflexiva ($r//r$ para toda recta r de E_n), simétrica (si $r//s$, entonces $s//r$) y transitiva (si $r//s$ y $s//t$, entonces $r//t$) y, por ello, es una equivalencia. Cada clase de esta equivalencia está formada por todas las rectas que tienen una misma dirección; hay correspondencia biyectiva entre las clases y las direcciones.

OBSERVACION

Hagamos una interpretación intuitiva de la anterior definición en el caso de E_2. Sean r_1 y r_2 dos rectas paralelas de E_2 y sea O un punto exterior a E_2 desde

el que proyectamos los puntos de r_1 y los de r_2. Si en r_1 se toma un punto P variable que tiende a infinito, los rayos OP tienen como posición límite a la recta r paralela a r_1 por O. Deberá entenderse que el punto P_∞, del infinito de la recta r_1, es el de intersección de E_2 con r. Si ahora se toma un punto Q sobre r_2 y se le hace tender a infinito, el rayo variable OQ tenderá, igual que el OP, a confundirse con la recta r; por ello, el punto del infinito Q_∞ de la recta r_2 deberá ser, también, el de intersección de E_2 con r. Ambas rectas, por ser paralelas, resultan tener el mismo punto del infinito $P_\infty = Q_\infty$. Si $\bar{u} \neq \bar{o}$ es un vector de la dirección de r_1 y de r_2, a dicho punto del infinito se le ha denotado antes por $\bar{u}_\infty$.

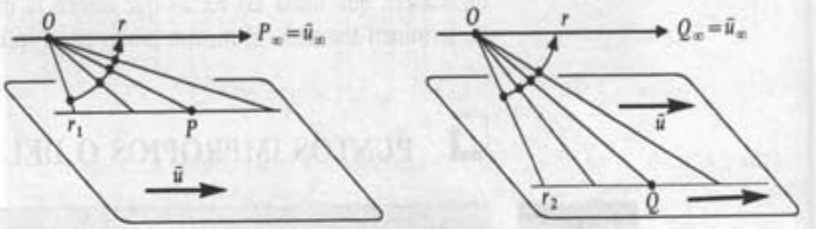


Figura 11.33.

ESPACIO AFIN AMPLIADO

[242]

Al añadir al espacio afín E_n ($n = 2$ o $n = 3$)(*) sus puntos impropios, se obtiene el espacio *afín ampliado* $\bar{E}_n$, o espacio afín proyectivizado. Si a una recta de E_n se le incorpora su punto del infinito, se obtiene una recta de $\bar{E}_n$ o recta ampliada. Si a un plano de E_n se le adicionan los puntos del infinito de todas sus rectas, se obtiene un plano de $\bar{E}_n$ o plano ampliado; se dice que los puntos del infinito de un plano de $\bar{E}_n$ forman su hiperplano del infinito (esto es: si $n = 2$, la recta del infinito; si $n = 3$, el plano del infinito).

Dos rectas paralelas de E_n, que tienen intersección vacía, al ampliarlas se cortan (en su punto del infinito). Dos planos paralelos de E_n, que tienen intersección vacía, al ampliarlos se cortan (a lo largo de su recta del infinito).

En $\bar{E}_2$, dos rectas cualesquiera se cortan en un punto (propio o impropio). En $\bar{E}_3$, dos planos cualesquiera se cortan a lo largo de una recta (propia o impropia).

(*) Cuanto aquí se dice, de E_n y de $\bar{E}_n$, es válido también para $n \in \mathbb{N}$ cualquiera.

OBSERVACION

Si en el espacio afín ampliado $\bar{E}_n$ ($n = 2$ o $n = 3$) no se hace distinción entre los puntos propios e impropios, considerando a todos iguales, no habrá posibi-

lidad de averiguar si dos rectas o planos dados son o no paralelos; esto es, habrá desaparecido la noción de paralelismo.

COORDENADAS HOMOGENEAS DE LOS PUNTOS DEL INFINITO

Al anterior concepto de puntos del infinito se le va a acompañar de un modo de asignarles coordenadas que, según tendremos ocasión de comprobar de aquí a poco, va a resultar muy útil y atinado, ya que conducirá a expresiones lineales en $\bar{E}_n$ de todo lo que era lineal en E_n, antes de ser ampliado.

[243]

Se llama *referencia cartesiana* del espacio afín ampliado $\bar{E}_n$ ($n=2$ o $n=3$)(*) a toda referencia cartesiana $(O; \bar{e}_1, ..., \bar{e}_n)$ de E_n. A los puntos $A_0 = O$ (origen) y $A_i = \bar{e}_{i\infty}$ (del infinito del eje i-ésimo; $i = 1, ..., n$) se les llama *puntos fundamentales* y al punto $A_u = O + (\bar{e}_1 + \cdots + \bar{e}_n)$ se le llama punto unidad de la referencia.

En $\bar{E}_n$ y respecto de la anterior referencia, se llaman *coordenadas homogéneas* de un punto del infinito $\bar{u}_\infty$ (del infinito de la dirección $\mathcal{V}(\bar{u})$, con $\bar{u} \neq \bar{o}$) a las $(0, \rho u_1, ..., \rho u_n)$, donde $(u_1, ..., u_n)$ son las coordenadas de $\bar{u} \in \mathbb{R}^n$ en la base $(\bar{e}_1, ..., \bar{e}_n)$ y $\rho \neq 0$ es arbitrario. Interesa destacar que las coordenadas $(0, u_1, ..., u_n)$ se pueden obtener como límite para $\lambda \to \infty$ de unas coordenadas homogéneas de $X = P + \lambda\bar{u}$ (con $P \in E_n$ fijo, cualquiera), punto genérico de una recta de dirección $\mathcal{V}(\bar{u})$.

(*) Cuanto aquí se dice es válido también para $n \in \mathbb{N}$ cualquiera.

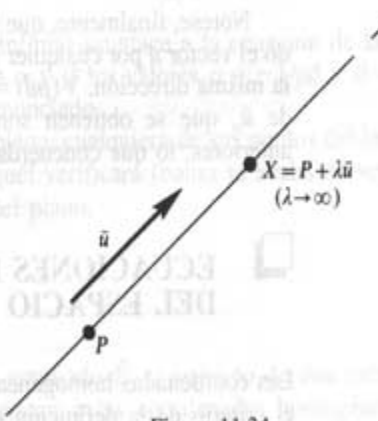


Figura 11.34.

COMPROBACION

Consideremos una recta cualquiera de las que tienen por dirección a $\mathcal{V}(\bar{u})$, esto es, una $r = P + \mathcal{V}(\bar{u})$ donde $P \in E_n$ se fija arbitrariamente. Si las coordenadas

cartesianas (absolutas) de P son $(p_1, ..., p_n)$, un punto genérico de r, que es el $X = P + \lambda\bar{u}$ para $\lambda \in \mathbb{R}$, tendrá por coordenadas cartesianas (absolutas) a los n números

$$x_1 = p_1 + \lambda u_1, ..., x_n = p_n + \lambda u_n \quad (\text{para } \lambda \in \mathbb{R}) \qquad [1]$$

Por tanto, el punto X, genérico de r, admite por coordenadas homogéneas a los $n + 1$ números siguientes, o a cualesquiera otros proporcionales a ellos:

$$x_0^* = 1, x_1^* = p_1 + \lambda u_1, ..., x_n^* = p_n + \lambda u_n \quad (\text{para } \lambda \in \mathbb{R}) \qquad [2]$$

Si ahora se hace que el punto $X = P + \lambda\bar{u}$ «tienda a infinito», lo que se consigue haciendo que λ diverja (esto es, para $\lambda \to \infty$), nos encontramos con que, tanto las coordenadas cartesianas [1] de X, como el sistema de coordenadas homogéneas [2] tienden a valores infinitos; resulta, pues, que con este modo de hacer las cosas no se consigue asignar coordenadas al punto del infinito $\bar{u}_\infty$ de la recta $r = P + \mathcal{V}(\bar{u})$. Lo que sí resulta útil es recurrir a otro sistema de coordenadas homogéneas del punto variable $X = P + \lambda\bar{u}$, que será proporcional al sistema [2] anterior; nos referimos, al siguiente sistema de coordenadas homogéneas de X:

$$\frac{x_0^*}{\lambda} = \frac{1}{\lambda}, \quad \frac{x_1^*}{\lambda} = \frac{p_1}{\lambda} + u_1, \quad ..., \quad \frac{x_n^*}{\lambda} = \frac{p_n}{\lambda} + u_n \quad (\text{para } \lambda \neq 0) \qquad [3]$$

Al hacer ahora que λ tienda a infinito, las coordenadas homogéneas [3] del punto X tienen por límites a $0, u_1, ..., u_n$, que son justamente las coordenadas homogéneas que, por definición, se asignaron al punto del infinito $\bar{u}_\infty$ de la recta $r = P + \mathcal{V}(\bar{u})$.

Nótese, finalmente, que todo el proceso anterior se puede repetir sustituyendo el vector $\bar{u}$ por cualquier otro vector $\rho\bar{u}$, con $\rho \neq 0$, ya que ambos determinan la misma dirección, $\mathcal{V}(\rho\bar{u}) = \mathcal{V}(\bar{u})$. Al hacerlo así, las coordenadas homogéneas de $\bar{u}_\infty$ que se obtienen son $(0, \rho u_1, ..., \rho u_n)$, que son proporcionales a las anteriores, lo que concuerda con la definición de coordenadas homogéneas.

ECUACIONES DE LAS RECTAS Y LOS PLANOS DEL ESPACIO AMPLIADO

Las coordenadas homogéneas que se les asigna a los puntos del infinito, según el criterio de la definición precedente [243], son aquellas que permiten tratar a dichos puntos del mismo modo que a los puntos propios de $\bar{E}_n$. Así, al considerar los puntos de infinito de una recta o de un plano, resulta que las coordenadas de éstos verifican a las ecuaciones de aquéllos, las cuales dan, pues, acogida a los puntos del infinito y, de este modo, son también ecuaciones de las respectivas rectas y planos del espacio ampliado. Veamos que las cosas son, en efecto, como se ha dicho:

[244]

En el espacio afín E_n ($n=2$ o $n=3$)(*) y respecto de cierta referencia cartesiana, considérense la recta $r(P, Q)$, que pasa por los puntos P y Q, y el plano $\pi(P, Q, R)$, que pasa por los puntos P, Q y R. Supongamos que las coordenadas de estos puntos son $P(\rho p_i^*)$, $Q(\rho q_i^*)$ y $R(\rho r_i^*)$ y que, por ello, las ecuaciones paramétricas (en coordenadas homogéneas) de la recta y el plano son, para $i=0, 1, ..., n$ (véase [239]):

$$r(P, Q): \rho x_i^* = \alpha p_i^* + \beta q_i^*$$

y

$$\pi(P, Q, r): \rho x_i^* = \alpha p_i^* + \beta q_i^* + \gamma r_i^*$$

Pues bien, estas ecuaciones lo son también de sus ampliaciones a $\bar{E}_n$, esto es, se verifican para sus puntos del infinito, los cuales se obtienen para los valores de (α, β) o de (α, β, γ) que conducen a $x_0^* = 0$.

(*) Cuanto aquí se dice es válido también para $n \in \mathbb{N}$ cualquiera.

COMPROBACION

Consideremos primero el caso de la recta. Las coordenadas cartesianas (absolutas) del vector $\overrightarrow{PQ}$ y las coordenadas homogéneas del punto del infinito $\overrightarrow{PQ}_\infty$ de la recta son:

$$\overrightarrow{PQ}\left(\frac{q_1^*}{q_0^*}-\frac{p_1^*}{p_0^*}, ..., \frac{q_n^*}{q_0^*}-\frac{p_n^*}{p_0^*}\right) \quad \text{y} \quad \overrightarrow{PQ}_\infty\left(0, \frac{q_1^*}{q_0^*}-\frac{p_1^*}{p_0^*}, ..., \frac{q_n^*}{q_0^*}-\frac{p_n^*}{p_0^*}\right)$$

Este punto del infinito satisface a la ecuación de la recta $r(P, Q)$ al tomar para los parámetros α y β los valores $\alpha = -1/p_0^*$ y $\beta = 1/q_0^*$, lo que confirma la afirmación del enunciado.

En el caso del plano, cualquiera de sus puntos del infinito lo es de una recta del plano, luego aquél verificará (como se acaba de ver) a la ecuación de ésta y, por tanto, a la del plano.

EJERCICIO

En el espacio afín ampliado $\bar{E}_3$ y respecto de una cierta referencia cartesiana (x_0^*, x_1^*, x_2^*, x_3^* denotan a las coordenadas homogéneas de un punto X), se considera el plano π que admite por ecuaciones homogéneas paramétricas a las:

$$x_0^* = \alpha - 2\beta - \gamma, \quad x_1^* = 2\alpha + \beta + \gamma, \quad x_2^* = -\alpha - \beta + 2\gamma, \quad x_3^* = 3\alpha + 2\beta - \gamma$$

Hallar ecuaciones paramétricas del lugar geométrico que forman los puntos del infinito de π (a este lugar geométrico se le llamará, en breve, recta del infinito).

RESOLUCION

Los puntos del infinito de π son aquellos que tienen $x_0^* = 0$, es decir, se obtienen para aquellos valores de los parámetros α, β y γ tales que $\alpha - 2\beta - \gamma = 0$. Por tanto sustituyendo γ por $\gamma = \alpha - 2\beta$, en las ecuaciones paramétricas de π, se obtiene el lugar geométrico pedido, que es, pues:

$$x_0^* = 0, \quad x_1^* = 3\alpha - \beta, \quad x_2^* = \alpha - 5\beta, \quad x_4^* = 2\alpha + 4\beta$$

RECTA DEL INFINITO (DE $\bar{E}_2$)Y PLANO DEL INFINITO (DE $\bar{E}_3$)

[245]

El conjunto de los puntos del infinito de $\bar{E}_2$ admite por ecuación a $x_0^* = 0$ (en cualquier referencia cartesiana de coordenadas homogéneas x_0^*, x_1^*, x_2^*) y se le llama *recta del infinito*. Toda recta de $\bar{E}_2$ (propia o del infinito) admite ecuación del tipo $a_0x_0^* + a_1x_1^* + a_2x_2^* = 0$ para ciertos a_0, a_1, a_2 tales que $(a_0, a_1, a_2) \neq (0, 0, 0)$ y recíprocamente.

El conjunto de los puntos del infinito de $\bar{E}_3$ admite por ecuación a $x_0^* = 0$ (en cualquier referencia cartesiana de coordenadas homogéneas x_0^*, x_1^*, x_2^*, x_3^*) y se le llama *plano del infinito*. Todo plano de $\bar{E}_3$ (propio o del infinito) admite ecuación del tipo $a_0x_0^* + a_1x_1^* + a_2x_2^* + a_3x_3^* = 0$ para ciertos a_0, a_1, a_2, a_3 tales que $(a_0, a_1, a_2, a_3) \neq (0, 0, 0, 0)$ y recíprocamente.

OBSERVACION

De acuerdo con la definición de coordenadas homogéneas, tanto para los puntos E_n (véase [237]) como para los puntos impropios (véase [243]), resulta que: respecto de una referencia cartesiana de $\bar{E}_n$, todo punto $X \in \bar{E}_n$ tiene por sistema de coordenadas homogéneas a un vector $\bar{x} = (x_0^*, x_1^*, ..., x_n^*)$ no nulo de $\mathbb{R}^{n+1}$ o a cualquiera de sus proporcionales $\rho\bar{x}$, con $\rho \neq 0$; si X es un punto propio, entonces $x_0^* \neq 0$, y si X es del infinito, entonces $x_0^* = 0$. Recíprocamente, un vector cualquiera $\bar{x} = (x_0^*, x_1^*, ..., x_n^*) \neq \bar{o}$ de $\mathbb{R}^{n+1}$ y todos sus proporcionales, $\rho\bar{x}$ con $\rho \neq 0$, son coordenadas homogéneas, en la referencia dada, de un cierto punto $X \in \bar{E}_n$; si $x_0^* \neq 0$, entonces X es un punto propio, y si $x_0^* = 0$, entonces X es un punto del infinito.

Así, pues, recurriendo a una referencia cartesiana de $\bar{E}_n$, se ha establecido una biyección entre los puntos $X \in \bar{E}_n$ y los conjuntos del tipo $\{\rho\bar{x}/\rho \in \mathbb{R}, \rho \neq 0\}$, donde $\bar{x}$ es un vector no nulo de $\mathbb{R}^{n+1}$.

ESPACIO PROYECTIVO ORDINARIO

La observación precedente nos lleva a concebir los puntos de $\bar{E}_n$ (propios o impropios) como si fueran subespacios unidimensionales de $\mathbb{R}^{n+1}$. Esta idea conduce a buen puerto, como se verá a continuación:

[246]

Se llama *espacio proyectivo* (ordinario) de dimensión n (para $n = 2$ o $n = 3$)(*) al conjunto $\mathcal{P}_n$ de los subespacios unidimensionales de $\mathbb{R}^{n+1}$, que se llaman puntos de $\mathcal{P}_n$; esto es:

$$\mathcal{P}_n = \{X = [\bar{x}] / \bar{x} \in \mathbb{R}^{n+1} - O\}$$

donde

$$[\bar{x}] = \mathcal{V}(\bar{x}) = \{\rho\bar{x} / \rho \in \mathbb{R}\}^{(**)}$$

Se llama *recta*, de $\mathcal{P}_n$, que pasa por dos puntos distintos $P = [\bar{p}]$ y $Q = [\bar{q}]$ al conjunto de los puntos $X = [\alpha\bar{p} + \beta\bar{q}]$ para $(\alpha, \beta) \in \mathbb{R}^2 - O$. Se llama *plano*, de $\mathcal{P}_n$, que pasa por los puntos no alineados $P = [\bar{p}]$, $Q = [\bar{q}]$ y $R = [\bar{r}]$ al conjunto de los puntos $X = [\alpha\bar{p} + \beta\bar{q} + \gamma\bar{r}]$ para $(\alpha, \beta, \gamma) \in \mathbb{R}^3 - O$.

Los espacios $\bar{E}_n$ (afín ampliado) y $\mathcal{P}_n$ (proyectivo) son isomorfos (identificables); esto es: Dada una referencia cartesiana en $\bar{E}_n$, que a cada punto $X \in \bar{E}_n$ le asigna unas coordenadas homogéneas $\rho\bar{x} = \rho(x_0^*, x_1^*, \ldots, x_n^*) \neq \bar{o}$ (donde $\rho \neq 0$ es cualquiera), la correspondencia $X \longmapsto [\bar{x}]$ es una biyección, de $\bar{E}_n$ en $\mathcal{P}_n$, que transforma rectas en rectas y planos en planos.

(*) Cuanto aquí se dice es válido para $n \in \mathbb{N}$ cualquiera.

(**) La notación $[\bar{x}]$ se podía omitir y utilizar la $\mathcal{V}(\bar{x})$, de siempre; no se hace así al objeto de abreviar.

COMPROBACION

En la «observación» precedente se ha justificado ya que la correspondencia $X \longmapsto [\bar{x}]$ es una biyección, de $\bar{E}_n$ en $\mathcal{P}_n$. Obsérvese que las rectas y los planos de $\mathcal{P}_n$ se han definido de manera que, en la anterior correspondencia, las rectas y los planos de $\bar{E}_n$ se transformen en las rectas y los planos de $\mathcal{P}_n$, según se desprende de la forma que toman las ecuaciones de unas y otros en $\bar{E}_n$ (véase [244]).

CAPITULO

12

Espacios puntuales o afines (caso general)

Si hacemos memoria, recordaremos que al principio, cuando empezábamos a hablar de vectores, echábamos una ojeada a lo que ocurría en la geometría elemental, donde un vector era un segmento orientado $\overrightarrow{AB}$, identificable con un par de puntos (A, B), que eran su origen y su extremo. Esta relación, que equipara un par de puntos con un vector, va a ser ahora nuestro punto de arranque para, conocidos ya suficientemente los vectores, iniciarnos en la geometría.

Así, pues, nuestro modo de abordar estas cuestiones altera lo que podíamos llamar el «orden natural», en el cual se pide que primero se estudien las cosas de los puntos y, apoyándose en ello, se analice después a los vectores. Aquí hemos utilizado el camino contrario, como es normal hoy día: se introdujeron los vectores de modo axiomático (fijándonos, claro está, en los vectores geométricos, que nos sirvieron de modelo), se estudiaron los espacios vectoriales y, recurriendo a todo esto, vamos a ocuparnos ahora de los puntos y demás elementos geométricos. El espacio que forman los puntos, que se llama *espacio afín*, está así asociado al espacio, vectorial, de los correspondientes vectores, que se llaman vectores libres.

Los problemas geométricos de incidencia, intersecciones y paralelismo, entre elementos del espacio afín, se abordarán con la ayuda de las propiedades relativas a la linealidad de los vectores libres. Pero, además de dichos problemas, llamados propiamente afines, están los problemas de tipo métrico, esto es, los relacionados con el cálculo de distancias, ángulos, áreas o volúmenes. Para resolver los problemas métricos vamos a necesitar del producto escalar, entre vectores libres, es decir, éstos deben formar un espacio vectorial euclídeo. El espacio afín asociado a un espacio vectorial euclídeo se llamará espacio afín euclídeo.

α ESPACIOS AFIN Y AFIN EUCLIDEO

Para abordar el estudio de la geometría, como el de cualquier otra teoría matemática, hemos de empezar diciendo cuál es la xiomática sobre la que nos apoyaremos. Tiene que quedar claro, desde el principio, cuáles son las propiedades que admitimos (porque sí; mejor dicho, para que las cosas funcionen como lo hacen en geometría elemental); todas las demás propiedades geométricas se irán deduciendo lógicamente a partir de aquéllas.

12.1. LOS ESPACIOS PUNTUALES

Con la siguiente definición se obliga a que cada pareja de puntos determine un vector, único, y recíprocamente, al modo de la geometría elemental:

PRIMERA DEFINICION DE ESPACIO AFIN

[247]

Un conjunto $E \neq \emptyset$, a cuyos elementos se les llama *puntos*, se dice que es un *espacio afín* asociado a un espacio vectorial real dado $V^{(*)}$, si se dispone de una aplicación

$$E \times E \to V, \quad (P, Q) \mapsto \vec{v} = \overrightarrow{PQ} \qquad [1]$$

que a cada par de puntos (P, Q) le asigna un vector $\vec{v}$ que se denota por $\overrightarrow{PQ}$, de manera que se verifiquen las dos exigencias siguientes:

1. Para cualesquiera que sean los puntos $P, Q, R \in E$, se verifica que

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR} \quad \text{(relación de Chasles)}$$

2. Para cada punto $P \in E$ y cada vector $\vec{v} \in V$, existe un punto $Q \in E$, y sólo uno, tal que $\vec{v} = \overrightarrow{PQ}$.

A los elementos de V se les llama *vectores libres*. Se dice que P y Q son, respectivamente, el *origen* y el *extremo* del vector $\overrightarrow{PQ}$.

(*) V podría ser (mientras que no se consideren cuestiones euclídeas) un espacio vectorial sobre un cuerpo K cualquiera.

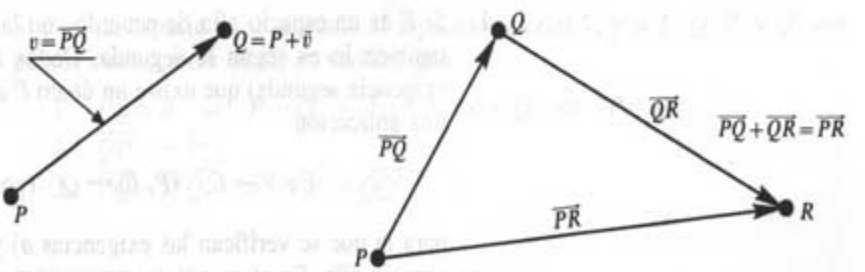


Figura 12.1.

OBSERVACION

Al referirnos a un espacio afín, diremos «sea el espacio afín E», cuando debiéramos mencionar también al espacio vectorial V asociado a E y a la aplicación $(P, Q) \mapsto \overrightarrow{PQ}$. Este abuso de lenguaje no creará dificultades; no

obstante, conviene no perder de vista que, según la definición precedente, el espacio afín que en ella se considera es la terna (E, V, Ψ), donde Ψ es la aplicación $\Psi: E \times E \to V$ que interviene en la anterior definición (véase [247], [1]).

SEGUNDA DEFINICION DE ESPACIO AFIN

Para expresar que $\vec{v} = \overrightarrow{PQ}$, también se suele poner $Q = P + \vec{v}$ (suma del punto P y del vector $\vec{v}$); con esta notación, el concepto de espacio vectorial puede expresarse también como sigue:

[248]

Un conjunto $E \neq \emptyset$, a cuyos elementos se les llama *puntos*, es un *espacio afín* asociado al espacio vectorial real V(*), si se dispone de una aplicación

$$E \times V \to E, \quad (P, \vec{v}) \mapsto Q = P + \vec{v}$$

que a cada punto P y cada vector $\vec{v}$ les asigna un punto Q, que se denota por $P + \vec{v}$ y se llama suma del punto P y del vector $\vec{v}$, de manera que se verifiquen las dos exigencias siguientes:

a) Para cada $P \in E$ y cada $Q \in E$, existe un vector $\vec{v} \in V$, y uno sólo, tal que $P + \vec{v} = Q$.

b) Para cualesquiera $P \in E$ y $\vec{u}, \vec{v} \in V$, se verifica que

$$P + (\vec{u} + \vec{v}) = (P + \vec{u}) + \vec{v}$$

(*) V podría ser (mientras no se consideren cuestiones euclídeas) un espacio vectorial sobre un cuerpo K cualquiera.

COMPROBACION

1. Si E es un espacio afín de acuerdo con la primera definición, veamos que también lo es según la segunda. Dados $P \in E$ y $\vec{v} \in V$, sabemos (por la exigencia segunda) que existe un único $P \in V$ tal que $\vec{v} = \overrightarrow{PQ}$, lo que define una aplicación

$$E \times V \to E, \quad (P, \vec{v}) \mapsto Q \quad \text{(pondremos } Q = P + \vec{v}\text{)}$$

para la que se verifican las exigencias *a*) y b) de la segunda definición de espacio afín. En efecto: la exigencia *a*) se verifica ya que $(P, Q) \mapsto \vec{v} = \overrightarrow{PQ}$ es aplicación (de acuerdo con la primera definición) y la exigencia *b*) es la exigencia segunda (de la primera definición), expresada de otro modo.

2. Si E es espacio afín de acuerdo con la segunda definición, veamos que también lo es según la primera. Dados $P, Q \in E$, sabemos (por la exigencia *a*)) que existe un único $\vec{v}$ tal que $P + \vec{v} = Q$, lo que define una aplicación

$$E \times E \to V, \quad (P, Q) \mapsto \vec{v} \quad \text{(pondremos } \vec{v} = \overrightarrow{PQ}\text{)}$$

para la que se verifican las exigencias primera y segunda (de la primera definición). En efecto: la exigencia primera no es otra cosa que la exigencia *a*) (de la segunda definición) expresada con otra notación y la exigencia segunda se verifica ya que $(P, \bar{v}) \mapsto Q = P + \bar{v}$ es aplicación (según reza en la segunda definición).

EJEMPLOS

1. Es obligado citar, en primer lugar, al plano y al espacio tridimensional ordinarios, los de la geometría elemental, que nos han servido de modelo para la anterior definición.
2. Sea $AX = B$ una ecuación matricial, en la que $A \in \mathcal{M}_{m\times n}$ y $B \in \mathcal{M}_{m\times 1}$ son matrices dadas y la incógnita es $X \in \mathcal{M}_{n\times 1}$. Si el sistema tiene alguna solución, el conjunto de todas ellas, E, es un espacio afín asociado al espacio vectorial V que forman las soluciones de la ecuación homogénea $AX = O$ (que es un subespacio de $\mathcal{M}_{n\times 1}$).
   Para comprobarlo, empecemos recordando que, según sabemos, si X_p y X_q son dos soluciones de la ecuación dada, entonces $X_v = X_q - X_p$ es solución de la homogénea y que, si X_p es solución de la ecuación dada y X_v es solución de la ecuación homogénea, entonces $X_q = X_p + X_v$ es solución de la ecuación dada. A la vista de lo dicho, la aplicación, de $E \times E$ en V, que se requiere (véase [247], [1]) para que E sea espacio afín será aquí la

$$E\times E \to V, \quad (X_p, X_q) \mapsto X_v = X_q - X_p$$

3. Un espacio vectorial V puede ser considerado como espacio afín (asociado a sí mismo). Para ello, basta con tomar para «suma de punto más vector» (véase [248]) la suma de vectores de V; para ella se verifican, obviamente, las exigencias *a*) y *b*) de la definición de espacio afín.

PRIMERAS PROPIEDADES

[249]

Si E es un espacio afín, asociado al espacio vectorial V, y si P, Q, P' y Q' son puntos cualesquiera de E, se verifica que:

1. $\overrightarrow{PQ} = \bar{o} \Leftrightarrow P = Q$ (o bien $P + \bar{o} = Q \Leftrightarrow P = Q$)
2. $\overrightarrow{QP} = -\overrightarrow{PQ}$
3. $\overrightarrow{PQ} = \overrightarrow{P'Q'} \Rightarrow \overrightarrow{PP'} = \overrightarrow{QQ'}$

DEMOSTRACION

1. De acuerdo con la exigencia primera de la definición [247], es $\overrightarrow{PP} + \overrightarrow{PP} = \overrightarrow{PP}$, de donde se infiere que $\overrightarrow{PP} = \bar{o}$. Recíprocamente, si $\overrightarrow{PQ} = \bar{o}$, como $\overrightarrow{PP} = \bar{o}$, resulta que, según la exigencia segunda de la definición [247], ha de ser $P = Q$.
2. Como (según la exigencia primera) es $\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = \bar{o}$, resulta que $\overrightarrow{QP} = -\overrightarrow{PQ}$.

3. Por verificarse que $\overrightarrow{PP'} = \overrightarrow{PQ} + \overrightarrow{QP'}$ y que $\overrightarrow{QQ'} = \overrightarrow{QP'} + \overrightarrow{P'Q'}$, restando miembro a miembro estas dos igualdades, se obtiene que

$$\overrightarrow{PP'} - \overrightarrow{QQ'} = \overrightarrow{PQ} - \overrightarrow{P'Q'}$$

luego la relación $\overrightarrow{PP'} = \overrightarrow{QQ'}$ equivale a la $\overrightarrow{PQ} = \overrightarrow{P'Q'}$

EJERCICIO

Sea E un espacio afín, asociado al espacio vectorial V. Se llama *traslación* de vector $\bar{v} \in V$ (dado) a la aplicación $T_{\bar{v}}$ siguiente:

$$T_{\bar{v}}: E \to E, \quad P \mapsto T_{\bar{v}}(P) = P + \bar{v}$$

Pruébese que:

1. $T_{\bar{v}}: E \to E$ es una biyección, para cualquiera que sea $\bar{v} \in V$.
2. $\mathcal{T} = \{T_{\bar{v}}: E \to E / \bar{v} \in V\}$ es un grupo, respecto de la composición de aplicación, que es isomorfo al grupo $(V, +)$ de los vectores libres.

RESOLUCION

1. La traslación $T_{\bar{v}}$ es una aplicación inyectiva ya que

$$T_{\bar{v}}(P) = T_{\bar{v}}(Q) \Rightarrow P + \bar{v} = Q + \bar{v} \Rightarrow P = Q$$

También es sobreyectiva ya que, dado $P' \in E$, existe $P \in E$ tal que $T_{\bar{v}}(P) = P'$; en efecto, tomando $P = P' + (-\bar{v})$, será:

$$T_{\bar{v}}(P) = P + \bar{v} = [P' + (-\bar{v})] + \bar{v} = P' + [(-\bar{v}) + \bar{v}] = P' + \bar{o} = P'$$

2. Veamos que $\mathcal{T}$ es un subgrupo del grupo de las biyecciones de E en E. Así ocurre ya que $\mathcal{T}$ no es vacío y se verifica que, dadas $T_{\bar{u}}, T_{\bar{v}} \in \mathcal{T}$, se verifica que $T_{\bar{u}} \circ T_{\bar{v}}$ y $T_{\bar{v}}^{-1}$ también son traslaciones; más exactamente $T_{\bar{u}} \circ T_{\bar{v}} = T_{\bar{u}+\bar{v}}$ y $T_{\bar{u}}^{-1} = T_{-\bar{u}}$, con lo que, además, el grupo $\mathcal{T}$ es isomorfo al grupo aditivo V, de los vectores libres. En efecto, para cualquiera que sea el punto $P \in E$, se tiene:

- $(T_{\bar{u}} \circ T_{\bar{v}})(P) = T_{\bar{u}}[T_{\bar{v}}(P)] = T_{\bar{u}}(P + \bar{v}) = (P + \bar{v}) + \bar{u} = P + (\bar{u} + \bar{v}) = T_{\bar{u}+\bar{v}}(P)$
- $T_u^{-1}(P) = P' \Leftrightarrow T_u(P') = P \Leftrightarrow P' + \bar{u} = P \Leftrightarrow P - \bar{u} = P' \Leftrightarrow T_{-\bar{u}}(P) = P'$

ESPACIO AFIN EUCLIDEO

Para que en los espacios afines tengan cabida los conceptos de distancia, ángulo, área y volumen (conceptos métricos), se precisa del producto escalar entre los vectores libres. Cuando se dispone de él, se dice que el espacio afín es euclídeo:

[250]

Sea E un espacio afín asociado a un espacio vectorial real V. Se dice que E es un *espacio afín euclídeo* (o espacio puntual euclídeo) si en su espacio vectorial asociado, V, hay definido un producto escalar, es decir, si V es un espacio vectorial euclídeo.

OBSERVACION

Según se dijo anteriormente, echando mano del producto escalar entre los vectores libres, vamos a poder hablar de distancias y ángulos entre los distintos elementos geométricos del espacio afín E. De la distancia vamos a ocuparnos inmediatamente. Respecto de los ángulos nos limitaremos, por el momento, a hacer el siguiente recordatorio:

Dado un ángulo $P\widehat{Q}R$ (donde $P, Q, R \in E$ y siendo $P \neq Q$ y $R \neq Q$), su coseno valdrá

$$\cos P\widehat{Q}R = \cos(\overrightarrow{PQ}, \overrightarrow{QR}) = \frac{\overrightarrow{QP} \cdot \overrightarrow{QR}}{\|\overrightarrow{QP}\| \, \|\overrightarrow{QR}\|}$$

En particular, si la medida de este ángulo es 0, π o $\pi/2$, entonces los vectores $\overrightarrow{QP}$ y $\overrightarrow{QR}$ son, respectivamente, de igual dirección y sentido, de la misma dirección y de sentido opuesto u ortogonales (perpendiculares).

DISTANCIA EUCLIDEA

[251]

Sea E un espacio afín euclídeo. Se llama *distancia* del punto $P \in E$ al punto $Q \in E$ a la norma(*) del vector $\overrightarrow{PQ}$, esto es, al número real:

$$d(P, Q) = \|\overrightarrow{PQ}\|$$

Para cualesquiera que sean los puntos $P, Q, R \in E$, se verifica que:

1. $d(P, Q) > 0$ para $P \neq Q$; además $d(P, P) = 0$.
2. $d(P, Q) = d(Q, P)$.
3. $d(P, R) \leq d(P, Q) + d(Q, R)$ (desigualdad triangular).

(*) Nos referimos, naturalmente, a la norma $\| \|$ del producto escalar; esto es, $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$ para todo vector libre $\vec{u}$.

DEMOSTRACION

Echando mano de las propiedades de la norma (véase [137]), se obtiene:

1. Si $P \neq Q$, es $d(P, Q) = \|\overrightarrow{PQ}\| > 0$, pues $\overrightarrow{PQ} \neq \bar{o}$

   $d(P, P) = \|\overrightarrow{PP}\| = \|\bar{o}\| = 0$

2. $d(P, Q) = \|\overrightarrow{PQ}\| = \|-\overrightarrow{QP}\| = |-1|\,\|\overrightarrow{QP}\| = \|\overrightarrow{QP}\| = d(Q, P)$
3. $d(P, R) = \|\overrightarrow{PR}\| = \|\overrightarrow{PQ} + \overrightarrow{QR}\| \leq \|\overrightarrow{PQ}\| + \|\overrightarrow{PQ}\| = d(P, Q) + d(Q, R)$

OBSERVACION (ESPACIO METRICO)

Las tres propiedades de la distancia euclídea, que acabamos de analizar en [251], se toman como axiomas cuando se generaliza el concepto de distancia. Esta generalización es como sigue:

Se llama *distancia* entre elementos (puntos) de un conjunto C a toda aplicación $d: C \times C \to \mathbb{R}$ tal que, para cualesquiera $x, y, z \in C$, se verifica:

1. $d(x, y) > 0$ si $x \neq y$; además $d(x, x) = 0$
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leq d(x, y) + d(y, z)$

Se llama *espacio métrico* a un conjunto C dotado de una distancia d; tal espacio se denota poniendo (C, d). Los espacios afines euclídeos son, pues, espacios métricos.

EJERCICIO

[252] Sea E un espacio afín euclídeo, asociado al espacio vectorial euclídeo V. Para cualesquiera que sean los puntos $P, Q, R \in E$ y el vector $\bar{u} \in V$, se verifica que:

1. $d(P, R) \geq |d(P, Q) - d(R, Q)|$
2. $d(P + \bar{u}, Q + \bar{u}) = d(P, Q)$

RESOLUCION

1. De acuerdo con la propiedad triangular (véase [251], 3), se verifica que

   $$\left.\begin{array}{l} d(P, Q) \leq d(P, R) + d(R, Q), \text{ luego } d(P, R) \geq d(P, Q) - d(R, Q) \\ d(R, Q) \leq d(R, P) + d(P, Q), \text{ luego } d(P, R) \geq d(R, Q) - d(P, Q) \end{array}\right\}$$

   y de estas dos últimas relaciones equivalen, obviamente, a la desigualdad que se deseaba demostrar.
2. Llamando $P' = P + \bar{u}$ y $Q' = Q + \bar{u}$, como $\overrightarrow{P'Q'} = \overrightarrow{PQ}$, se tiene:

   $$d(P + \bar{u}, Q + \bar{u}) = d(P', Q') = \|\overrightarrow{P'Q'}\| = \|\overrightarrow{PQ}\| = d(P, Q)$$

12.2. COORDENADAS CARTESIANAS

Nos vamos a referir aquí a los espacios afines y afines euclídeos de dimensión finita. En ellos es posible caracterizar los puntos mediante coordenadas, lo que va a permitir obtener ecuaciones de rectas, de planos y, en general, de variedades lineales cualesquiera. Para ellos procederemos guiados por lo que se hizo, para obtener coordenadas de los vectores, en los espacios vectoriales de dimensión finita, en los que se recurrió a las bases. Aquí echaremos mano de las *referencias cartesianas*, que no son otra cosa que bases de vectores libres a las que se les añade un *origen* (punto fijo del espacio); una tal referencia puede concebirse como la familia de los vectores fijos que resultan de «aplicar» en el origen los vectores libres de la base.

ESPACIO AFIN DE DIMENSION FINITA

[253]

Se dice que un espacio afín E tiene *dimensión* $n \in \mathbb{N}$ si su espacio vectorial asociado V tiene dimensión n; se pone $\dim E = n$. Si $n = 1$, se dice que E es una *recta afín*; si E tiene dimensión $n = 2$, se dice que E es un *plano afín*; si E tiene dimensión $n = 3$, se dice que E es un *espacio tridimensional afín*. Si $n = 0$, es $V = O$ y E se reduce a un punto.

Si E es un espacio afín de dimensión n, puede suponerse que sus vectores libres son los vectores de K^n (donde K es el cuerpo de escalares), ya que el espacio vectorial V asociada a E, por tener dimensión n, es isomorfo a K^n.

REFERENCIA CARTESIANA. COORDENADAS

Si recurrimos a un punto fijo O del espacio afín E, la correspondencia que a cada punto $X \in E$ le atribuye el vector $\overrightarrow{OX}$ es una biyección. Resulta entonces que, si E tiene dimensión finita y elegida una base de los vectores libres, las coordenadas del vector $\overrightarrow{OX}$ determinan de manera inequívoca al punto X. De ahí que se defina:

[254]

Sea E un espacio afín de dimensión finita n, asociado al espacio vectorial V. Se dice que $(O; \vec{e}_1, \vec{e}_2, ..., \vec{e}_n)$ es una *referencia cartesiana* del espacio afín si O es un punto de E, que se llama *origen*, y $(\vec{e}_1, \vec{e}_2, ..., \vec{e}_n)$ es una base del espacio V de los vectores libres.

Se llaman *coordenadas cartesianas* de un punto $X \in E$ en la referencia $(O; \vec{e}_1, \vec{e}_2, ..., \vec{e}_n)$ a las coordenadas $(x_1, x_2, ..., x_n)$ del vector $\overrightarrow{OX}$ en la base $(\vec{e}_1, \vec{e}_2, ..., \vec{e}_n)$ de V; se pone $X(x_i)$, donde $i = 1, 2, ..., n$.

Si $P(p_i)$ y $Q(q_i)$, entonces las coordenadas del vector $\overrightarrow{PQ}$ son los escalares $q_i - p_i$ para $i = 1, 2, ..., n$.

Antes de realizar las pertinentes comprobaciones, digamos que las rectas (que definiremos, en breve, de igual modo que ya se hizo en los casos $n = 2$ y $n = 3$) que pasan por O y tienen las direcciones de los vectores $\bar{e}_1$ (para $i = 1, 2, ..., n$) se llaman «ejes» de la referencia cartesiana $(O; \bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$. En ocasiones, en lugar de hablar de la referencia cartesiana, se habla de «los ejes de referencia», lo que no es ortogoxo ya que del conocimiento de los ejes, sin más, no se deduce cuáles son los vectores $\bar{e}_i$.

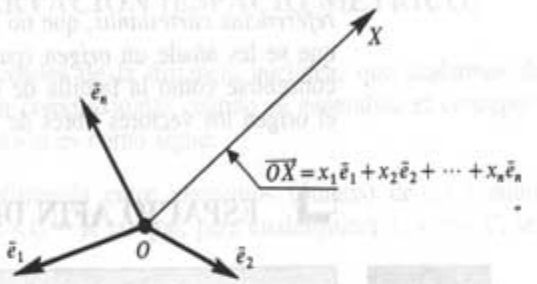


Figura 12.2.

COMPROBACIONES

Una vez fijada la referencia $(O, \bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$, a cada punto $X \in E$ le corresponde un solo vector $\overrightarrow{OX} \in V$, el cual tiene una única sucesión de n coordenadas $(x_1, x_2, ..., x_n)$, en la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$, que son los escalares tales que:

$$\overrightarrow{OX} = x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n$$

Recíprocamente, dada cualquier sucesión $(x_1, x_2, ..., x_n)$, de n escalares, el vector $\bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n$ define a un único punto $X \in E$ tal que $\bar{x} = \overrightarrow{OX}$; los escalares dados son las coordenadas del punto X.

Nótese, finalmente, que si P y Q son dos puntos de E cuyas coordenadas en la referencia dada son

$$P(p_1, p_2, ..., p_n) \quad \text{y} \quad Q(q_1, q_2, ..., q_n)$$

entonces

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \sum_{i=1}^{n} q_i\bar{e}_i - \sum_{i=1}^{n} p_i\bar{e}_i = \sum_{i=1}^{n} (q_i - p_i)\bar{e}_i$$

de donde resulta que las coordenadas del vector $\overrightarrow{PQ}$ en la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ son las $q_i - p_i$ (para $i = 1, 2, ..., n$).

CAMBIO DE COORDENADAS

Supongamos que se dispone de dos referencias distintas, en un mismo espacio afín; para cada punto del espacio habrá, pues, dos sistemas de coordenadas. Vamos a obtener la relación que liga a unas y otras coordenadas:

[255]

Sea E un espacio afín de dimensión finita n, asociado al espacio vectorial V, y sea $(O; \vec{e}_1, \vec{e}_2, ..., \vec{e}_n)$ una referencia de E; llamemos $(x_1, x_2, ..., x_n)$ a las coordenadas de un punto $X \in E$ en la citada referencia. Supóngase que $(C; \vec{u}_1, \vec{u}_2, ..., \vec{u}_n)$ es una nueva referencia de E; llamemos $(x'_1, x'_2, ..., x'_n)$ a las coordenadas del punto X en esta nueva referencia.

Si se conoce la segunda referencia en función de la primera, esto es, si se conocen los escalares c_i y q_{ij} que permiten poner:

$$\overrightarrow{OC} = \sum_{i=1}^{n} c_i \vec{e}_i \quad \text{y} \quad \vec{u}_j = \sum_{i=1}^{n} q_{ij} \vec{e}_i \quad (\text{para } j = 1, 2, ..., n)$$

entonces las coordenadas (x_i) y (x'_j) están relacionadas mediante:

$$x_i = c_i + \sum_{j=1}^{n} q_{ij} x'_j \quad \text{o} \quad X = C + QX'$$

donde(*)

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad X' = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix}, \quad C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \quad Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix}$$

La matriz Q es regular. Nótese que la relación matricial $X = C + QX'$ puede expresarse en la forma:

$$\begin{bmatrix} 1 \\ X \end{bmatrix} = \begin{bmatrix} 1 & O \\ C & Q \end{bmatrix} \begin{bmatrix} 1 \\ X' \end{bmatrix}$$

(*) Cada una de las letras X y C ha designado a un punto; ahora también va a representar a la matriz columna de las coordenadas de dicho punto (en la primera referencia). Ello no es un inconveniente pues, por el contexto de lo que se diga, no va a haber dudas sobre lo que representan en cada caso.

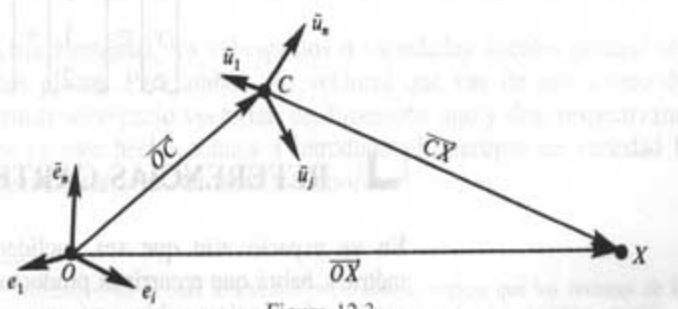


Figura 12.3.

COMPROBACION

Poniendo que $\overrightarrow{OX} = \overrightarrow{OC} + \overrightarrow{CX}$ y expresando estos vectores en función de las bases $(\bar{e}_i)$ y $(\bar{u}_j)$, que luego relacionamos entre sí, se obtiene que

$$\overrightarrow{OX} = \overrightarrow{OC} + \overrightarrow{CX} \Rightarrow \sum_{i=1}^{n} x_i \bar{e}_i = \sum_{i=1}^{n} c_i \bar{e}_i + \sum_{j=1}^{n} x'_j \bar{u}_j = \sum_{i=1}^{n} c_i \bar{e}_i +$$

$$+ \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} q_{ij} \bar{e}_i \right) = \sum_{i=1}^{n} \left[c_i + \sum_{j=1}^{n} x'_j q_{ij} \right] \bar{e}_i \Rightarrow x_i = c_i + \sum_{j=1}^{n} x'_j q_{ij}$$

Esta última relación, que es la que buscábamos, puede expresarse también, obviamente, de forma matricial poniendo $X = C + QX$. Nótese que Q es la matriz de cambio de coordenadas, en el espacio vectorial V de los vectores libres, ligado al cambio de la base $(\bar{e}_i)$ por la $(\bar{u}_j)$; por tanto (véase [069]) Q es una matriz regular.

EJERCICIO

En el espacio tridimensional afín y respecto de cierta referencia cartesiana $(O; \bar{e}_1, \bar{e}_2, \bar{e}_3)$, se considera el tetraedro $ABCD$, cuyos vértices tienen las siguientes coordenadas:

$$A(3, 1, -2), \quad B(2, 2, 0), \quad C(1, 0, -1) \quad \text{y} \quad D(4, 3, -2)$$

Si en la nueva referencia $(A; \overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD})$ las coordenadas de un punto P son (x', y', z'), hallar las coordenadas (x, y, z) de P en la primera referencia.

RESOLUCION

Como las coordenadas de $\overrightarrow{AB}$, $\overrightarrow{AC}$ y $\overrightarrow{AD}$ en la base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ son

$$\overrightarrow{AB}(-1, 1, 2), \quad \overrightarrow{AC}(-2, -1, 1) \quad \text{y} \quad \overrightarrow{AD}(1, 2, 0)$$

de acuerdo con el resultado anterior, será:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix} + \begin{bmatrix} -1 & -2 & 1 \\ 1 & -1 & 2 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

REFERENCIAS CARTESIANAS RECTANGULARES

En un espacio afín que sea euclídeo, para abordar los problemas de tipo métrico, habrá que recurrir al producto escalar y, para que la expresión de éste sea lo más simple posible, será entonces aconsejable utilizar bases de vectores

libres que sean ortonormales. Un sistema de referencia que tenga la base ortonormal se llamará rectangular(*):

[256]

En un espacio afín euclídeo E de dimensión finita n, asociado al espacio vectorial V, se dice que una referencia cartesiana $(O; \bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ es *rectanguloar* si la base $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_n)$ de V es ortonormal, esto es, si $\|\bar{e}_i\| = 1$ y $\bar{e}_i \cdot \bar{e}_j = 0$ para $i \neq j$ $(i, j = 1, 2, ..., n)$.

Si $(C; \bar{u}_1, \bar{u}_2, ..., \bar{u}_n)$ es otra referencia cartesiana, de E, que también es rectangular, entonces el correspondiente cambio de coordenadas, esto es, $X = C + QX'$ (véase [255]), tiene la matriz Q ortogonal.

Si, en una referencia rectangular, las coordenadas de dos puntos $X, Y \in E$ son $X(x_i)$ e $Y(y_i)$, para $i = 1, 2, ..., n$, entonces la distancia entre ellos vale:

$$d(X, Y) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2}$$

COMPROBACIONES

1. La matriz Q es la del cambio de coordenadas, en el espacio V de los vectores libres, que se produce al cambiar la base $(\bar{e}_i)$ por la base $(\bar{u}_j)$; como ambas bases son ortonormales, resulta que Q es ortogonal (véase [149]).
2. Si la referencia $(O; \bar{e}_i)$ es rectangular, o sea, si la base $(\bar{e}_i)$ de vectores libres es ortonormal, la matriz métrica del producto escalar es la unidad y, por ello, resulta que:

$$d(X, Y) = \|\overrightarrow{XY}\| = \left\| \sum_{i=1}^{n} (y_i - x_i)\bar{e}_i \right\| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + - (y_n - x_n)^2}$$

12.3. VARIEDADES LINEALES AFINES

En geometría elemental, los subespacios o variedades lineales (afines) son las rectas y los planos. Para ambos, los vectores que van de uno a otro de sus puntos forman subespacio vectorial, de dimensión uno y dos, respectivamente. Fijándonos en este hecho, vamos a introducir el concepto de variedad lineal afín, que generaliza a los de recta y plano.

(*) Este calificativo no es muy afortunado pues, aunque sugiere que los vectores de la base son ortogonales, no refleja el hecho de que dichos vectores han de ser, además, unitarios.

VARIEDADES LINEALES

[257]

Sea E un espacio afín, asociado al espacio vectorial V. Dados un punto $P \in E$ y un subespacio vectorial U de V, se llama *variedad lineal* (afín) de E que pasa por P y tiene la *dirección* U al siguiente conjunto de puntos de E, que denotaremos por $P + U$:

$$P + U = \{P + \bar{u}/\bar{u} \in U\} = \{X \in E/\overrightarrow{PX} \in U\}$$

Si dim $U = m$, se dice que $P + U$ tiene dimensión m; en particular, si $m = 1$ o $m = 2$, se dice que $P + U$ es una *recta* o un *plano*, respectivamente. Si dim $V = n$ y dim $U = n - 1$, se dice que $P + U$ es un *hiperplano*.

EJERCICIO

Sea $F \subset E$ un conjunto no vacío de puntos, del espacio afín E, tal que el conjunto $U = \{\overrightarrow{XY}/X, Y \in F\}$ es un subespacio vectorial, del espacio vectorial V de los vectores libres. Compruébese que F está incluido en una variedad lineal de E que tiene la dirección de V, pero que puede no coincidir con ella.

COMPROBACION

Como F no es vacío, existe $P \in F$. Comprobemos que $F \subset P + U$. Si X es un punto de F entonces $\overrightarrow{PX} \in U$, ya que P también es de F; por tanto, $P + \overrightarrow{PX}$ es un punto de $P + U$, es decir, X es de $P + U$, como había que comprobar.

Para concluir, pongamos un ejemplo en el que $P + U$ tenga algún punto más que F. Este es el caso, por ejemplo, de una recta $P + U$ (con dim $U = 1$) a la que se le quita un punto A; el conjunto $F = (P + U) - \{A\}$ no coincide con $P + U$ aun cuando, como resulta evidente, el conjunto $\{\overrightarrow{XY}/X, Y \in F\}$ es U.

IGUALDAD DE VARIEDADES LINEALES

Una variedad lineal, que queda determinada por uno de sus puntos y por un subespacio vectorial (la dirección), puede definirse también por cualquier otro de sus puntos; la dirección es única. Más exactamente:

[258]

Dadas dos variedades lineales $P_1 + U_1$ y $P_2 + U_2$, de un espacio afín E (asociado al espacio vectorial V), se verifica que

$$[P_1 + U_1 = P_2 + U_2] \Leftrightarrow [\overrightarrow{P_1P_2} \in U_1 = U_2]$$

DEMOSTRACION

1. Comprobaremos primeramente la implicación hacia la izquierda ($\Leftarrow$). Si $X \in P_1 + U_1$, entonces $\overrightarrow{XP_1} \in U_1$, luego $\overrightarrow{XP_1} \in U_2$ y, como $\overrightarrow{P_1P_2} \in U_2$, resulta que $\overrightarrow{XP_2} = \overrightarrow{XP_1} + \overrightarrow{P_1P_2}$ también pertenece a U_2, de donde se infiere que $X \in P_2 + U_2$. Análogamente se obtiene que si $X \in P_2 + U_2$ entonces $X \in P_1 + U_1$.
2. Veamos ahora que también se verifica la otra implicación ($\Rightarrow$). Como $P_1 \in P_1 + U_1$, se verifica que $P_1 \in P_2 + U_2$, luego $\overrightarrow{P_1P_2} \in U_2$; análogamente se comprueba que $\overrightarrow{P_1P_2} \in U_1$. Por otra parte, dado $\bar{u}_1 \in U_1$, como $P_1 + \bar{u}_1 \in P_1 + U_1$, resulta que $P_1 + \bar{u}_1 \in P_2 + U_2$ y, por tanto, es $P_1 + \bar{u}_1 = P_2 + \bar{u}_2$ para algún $\bar{u}_2 \in U_2$; de esta última igualdad se deduce que $\bar{u}_1 = \overrightarrow{P_1P_2} + \bar{u}_2$ y, como $\overrightarrow{P_1P_2}$ y $\bar{u}_2$ pertenecen ambos a U_2, se concluye que $\bar{u}_1 \in U_1$. Análogamente, si $\bar{u}_2 \in U_2$ se verificará que $\bar{u}_2 \in U_1$ y, por todo ello, ha de ser $U_1 = U_2$.

EJERCICIO

Si A y B son dos puntos de una variedad lineal, $P + U$, de un cierto espacio afín, entonces $\overrightarrow{AB}$ es un vector de la dirección de la variedad, esto es, $\overrightarrow{AB} \in U$.

COMPROBACION

Dado que A es un punto de $P + U$, es $P + U = A + U$. Por otra parte, como B es un punto de $A + U$, resulta finalmente que $\overrightarrow{AB}$ es un vector de U.

PARALELISMO

Recuérdese que, según sabemos de geometría elemental, dos rectas o dos planos son paralelos si tienen la misma dirección; una recta era paralela a un plano si la dirección de la recta estaba incluida en la del plano. Una generalización inmediata de ello conduce a la siguiente definición:

[259]

Sean $P_1 + U_1$ y $P_2 + U_2$ dos variedades lineales, de un espacio afín E (asociado al espacio vectorial V). Se dice que dichas variedades son *paralelas* si la dirección de una de ellas está incluida en la dirección de la otra o si, en particular, ambas direcciones coinciden; esto es, si $U_1 \subset U_2$, $U_2 \subset U_1$ o $U_1 = U_2$.

Si dos variedades lineales son paralelas, entonces o tienen intersección vacía o una de ellas está incluida en la otra.

DEMOSTRACION

Si $P_1 + U_1$ y $P_2 + U_2$ son paralelas y tienen intersección no vacía, recurriendo a un punto P que pertenezca a ambas, éstas se pueden expresar en la forma

$P+U_1$ y $P+U_2$. Como $U_1 \subset U_2$ o $U_2 \subset U_1$, resulta obvio que $P+U_1 \subset P+U_2$ o que $P+U_2 \subset P+U_1$, como había que comprobar.

ORTOGONALIDAD O PERPENDICULARIDAD

Dos rectas son perpendiculares, según sabemos de geometría elemental, cuando sus direcciones son subespacios vectoriales ortogonales. Otro tanto cabe decir acerca de la perpendicularidad de recta y plano, que acontece cuando las direcciones de recta y plano son subespacios vectoriales ortogonales. Sin embargo, la perpendicularidad entre planos se define de otra manera: dos planos son perpendiculares si son perpendiculares dos rectas perpendiculares a uno y otro plano. Respondiendo a todo ello, se da la siguiente definición para un caso general.

[260]

Sea P_1+U_1 y P_2+U_2 dos variedades lineales, de un espacio afín euclídeo E (asociado al espacio vectorial V). Si los subespacios vectoriales U_1 y U_2 son ortogonales (o sea, $\vec{u}_1 \cdot \vec{u}_2 = 0$ para cualesquiera $\vec{u}_1 \in U_1$ y $\vec{u}_2 \in U_2$), se dice que las variedades lineales son *perpendiculares* u *ortogonales*. Obsérvese que, si dim $E = n$, para que dichas variedades sean perpendiculares es necesario que dim U_1 + dim $U_2 \leqslant n$.

Si dim $E = n$ y dim U_1 + dim $U_2 > n$, se amplía la definición de perpendicularidad diciendo que P_1+U_1 y P_2+U_2 son perpendiculares (en sentido amplio) si los subespacios ortogonales de U_1 y U_2 (véase [144]) son ortogonales entre sí, esto es, si $U_1^{\perp} \perp U_2^{\perp}$.

12.4. VARIEDADES LINEALES DE DIMENSION FINITA

VARIEDAD LINEAL ENGENDRADA POR UN CONJUNTO DE PUNTOS

Según ya sabemos, de geometría elemental, la recta que pasa por dos puntos (distintos) es, de entre todas las variedades lineales que pasan por dichos puntos, la más pequeña de todas, esto es, la que está incluida en todas las demás variedades que pasan por dichos dos puntos (estas últimas variedades son planos y el propio espacio tridimensional); se dice que dicha recta es la variedad lineal que engendran los dos puntos. En general, cuando, en lugar de dos puntos, se considera un conjunto cualquiera de puntos, se engendra la siguiente variedad lineal:

[261]

Dado un conjunto de puntos, C, de un espacio afín E (asociado al espacio vectorial V), sea $P + U$ la variedad afín en la que P es un punto fijo de C y U es el subespacio engendrado por todos los vectores libres $\overrightarrow{XY}$, donde $X, Y \in C$. Se verifica que $P + U$ es la menor (menos amplia) de todas las variedades lineales de E que incluyen a C. Se dice que $P + U$ es la variedad afín engendrada por el conjunto de puntos C.

Si C el conjunto de puntos $C = \{P_0, P_1, ..., P_m\}$, entonces la variedad lineal engendrada por C es:

$$P_0 + V(\overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_m}) \qquad [1]$$

DEMOSTRACION

Obsérvese en primer lugar que C está incluido en la variedad $P + U$ ya que, si $Q \in C$ y como $\overrightarrow{PQ} \in U$, resulta que $Q = P + \overrightarrow{PQ}$ pertenece a $P + U$. Así, pues, todo lo que debemos comprobar es que, si F es una variedad lineal que incluye a C, entonces F incluye a $P + U$. Ahora bien, como $P \in C \subset F$, se puede poner $F = P + W$, donde W es un cierto subespacio de vectores libres; de ahí que lo que debemos comprobar es que $U \subset W$.

Si $\bar{u} \in U$, entonces $\bar{u}$ es una combinación lineal de los vectores $\overrightarrow{XY}$, para $X, Y \in C$, lo que expresaremos abreviadamente poniendo $\bar{u} = \Sigma \lambda \overrightarrow{XY}$ (donde λ representa a un escalar genérico). Ahora bien, como $X, Y \in C \subset P + W$, resulta que $\overrightarrow{XY} \in W$ (véase el ejercicio anterior) y, por ser W un espacio vectorial, se ha de verificar que $\Sigma \lambda \overrightarrow{XY} \in W$, es decir, que $\bar{u} \in W$, lo que prueba que $U \subset W$.

Finalmente, para comprobar que [1] es la variedad afín que engendran los puntos del conjunto $C = \{P_0, P_1, ..., P_n\}$, solo hay que ver que

$$\{\overrightarrow{P_0P_i}/P_i \in C\} \quad \text{y} \quad \{\overrightarrow{P_iP_j}/P_i, P_j \in C\}$$

engendran el mismo subespacio vectorial. Esto será cierto si cualquier vector $\overrightarrow{P_iP_j}$ del segundo conjunto resulta ser combinación lineal de vectores del primero, lo cual es evidente, ya que $\overrightarrow{P_iP_j} = \overrightarrow{P_0P_j} - \overrightarrow{P_0P_i}$.

VARIEDAD LINEAL DE DIMENSION FINITA. ECUACIONES PARAMETRICAS

Se llama dimensión de una variedad lineal $P + U$ a la dimensión del subespacio vectorial U; si dicha dimensión es p, recurriendo a una base de U, un punto genérico de $P + U$ se puede expresar en función de p parámetros (escalares) independientes, lo que permite obtener ecuaciones paramétricas (con p parámetros) de la variedad dada:

[262]

Se dice que una variedad lineal $P + U$, de un espacio afín E (asociado al espacio vectorial V), tiene *dimensión* finita p si la dirección U de la variedad es un subespacio vectorial de dimensión p. Si dim $U = p$ y si $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es una base de U, entonces la variedad $P + U$ está formada por los puntos $X \in E$ que se pueden poner en la forma (ecuación paramétrica vectorial):

$$X = P + \lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p \qquad [1]$$

($\lambda_1, \lambda_2, ..., \lambda_p$ escalares arbitrarios).

Si el espacio E tiene dimensión n y tomando una referencia cartesiana en él, si las coordenadas de X, P y $\bar{u}_j$ se denotan por $X(x_i)$, $P(p_i)$ y $\bar{u}_j(u_{ji})$ para $i = 1, 2, ..., n$ y $j = 1, 2, ..., p$, entonces los puntos de la variedad $P + U$ son aquellos X tales que:

$$x_i = p_i + \sum_{j=1}^{p} \lambda_j u_{ji} \quad (i = 1, 2, ..., n)$$

esto es:

$$\begin{cases} x_1 = p_1 + \lambda_1 u_{11} + \lambda_2 u_{21} + \cdots + \lambda_p u_{p1} \\ x_2 = p_2 + \lambda_1 u_{12} + \lambda_2 u_{22} + \cdots + \lambda_p u_{p2} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = p_n + \lambda_1 u_{1n} + \lambda_2 u_{2n} + \cdots + \lambda_p u_{pn} \end{cases} \qquad [2]$$

($\lambda_1, \lambda_2, ..., \lambda_p$ escalares arbitrarios, parámetros independientes).

COMPROBACION

Como $U = \mathcal{V}(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$, los puntos de $P + U$ son aquella X tales que

$$\overrightarrow{PX} = \lambda_1\bar{u}_1 + \lambda_2\bar{u}_2 + \cdots + \lambda_p\bar{u}_p \qquad [*]$$

para algunos escalares $\lambda_1, \lambda_2, ..., \lambda_p$. Esta relación es equivalente a la [1] del enunciado. Expresando que las coordenadas del primer y segundo miembros de [*] son iguales, se obtiene

$$x_i - p_i = \lambda_1 u_{1i} + \lambda_2 u_{2i} + \cdots + \lambda_p u_{pi} \quad (\text{para } i = 1, 2, ..., n)$$

que es equivalente a la relación [2].

Nótese que, como $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_p)$ es una base de U, resulta que cada punto X de $P + U$ queda definido por unos únicos valores de los parámetros $\lambda_1, \lambda_2, ..., \lambda_p$.

EJEMPLO

En un espacio afín real de dimensión 4 y respecto de una cierta referencia cartesiana, se consideran los puntos P, Q y R que tienen por coordenadas a:

$$P(1, 2, -1, 0), \quad Q(0, 3, 1, -2), \quad R(4, 0, -1, 1)$$

El plano que pasa por estos tres puntos (o plano que engendran dichos puntos), que es el conjunto de puntos X que se pueden poner:

$$X = P + \lambda\overrightarrow{PQ} + \mu\overrightarrow{PR}, \quad \text{para } \lambda, \mu \in \mathbb{R}$$

admite, pues, las siguientes ecuaciones paramétricas:

$$\begin{cases} x_1 = 1 - \lambda + 3\mu \\ x_2 = 2 + \lambda - 2\mu \\ x_3 = -1 + 2\lambda \\ x_4 = 0 - 2\lambda + \mu \end{cases} \quad (\lambda, \mu \in \mathbb{R})$$

Si se eliminan los parámetros λ y μ, se obtienen las siguientes ecuaciones (implícitas) del plano en cuestión:

$$\begin{cases} 2x_1 - 2x_3 - 3x_4 = 4 \\ 2x_2 + x_3 + 2x_4 = 3 \end{cases}$$

HIPERPLANOS

En un espacio afín E, se llaman hiperplanos a aquellas variedades lineales que no están incluidas en otra variedad más amplia, salvo en el propio espacio E. Si $\dim E = n$, la hiperplano son sus variedades lineales de dimensión $n - 1$:

[263]

En un espacio afín E de dimensión finita n, asociado al espacio vectorial V, se llaman *hiperplanos* a las variedades lineales de dimensión $n - 1$. Si $P + U$ es un hiperplano, donde $\dim U = n - 1$, y si $(\vec{u}_1, \vec{u}_2, ..., \vec{u}_{n-1})$ es una base de U, dicho hiperplano está formado por los puntos X tales que

$$\text{rang } \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_{n-1}, \overrightarrow{PX}\} = n - 1 \qquad [1]$$

Respecto de una referencia cartesiana, cualquier hiperplano admite ecuación de la forma

$$\begin{matrix} a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \\ \text{con } (a_1, a_2, ..., a_n) \neq (0, 0, ..., 0) \end{matrix} \quad \begin{pmatrix} \text{ecuación general} \\ \text{de un hiperplano} \end{pmatrix} \qquad [2]$$

donde $a_0, a_1, a_2, ..., a_n$ son escalares fijos y $(x_1, x_2, ..., x_n)$ son las coordenadas de un punto genérico del hiperplano. Recíprocamente, todos los puntos que verifican a una ecuación del tipo [2] forman un hiperplano de E.

COMPROBACIONES

El hiperplano $P + U$ está formado por los puntos X tales que $\overrightarrow{PX} \in U$; como $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_{n-1})$ es base de U, la condición $\overrightarrow{PX} \in U$ equivale a la anterior relación [1].

La pertenencia de X al hiperplano $P + U$, esto es, el hecho de que $\overrightarrow{PX}$ dependa linealmente de los vectores $\bar{u}_1, \bar{u}_2, ..., \bar{u}_{n-1}$, se puede expresar, recurriendo a las coordenadas, $X(x_i)$, $P(p_i)$ y $\bar{u}_j(u_{ji})$. diciendo que la columna de las coordenadas de $\overrightarrow{PX}$ depende linealmente de las columnas de coordenadas de los $n-1$ vectores $\bar{u}_j$, lo que equivale a:

$$\begin{vmatrix} u_{11} & \cdots & u_{n-11} & x_1 - p_1 \\ u_{12} & \cdots & u_{n-12} & x_2 - p_2 \\ & & & \\ & & & \\ u_{1n} & \cdots & u_{n-1n} & x_n - p_n \end{vmatrix} = 0$$

y desarrollando este determinante por los elementos de su última columna, se obtiene la siguiente ecuación del hiperplano:

$$a_1(x_1 - p_1) + a_2(x_2 - p_2) + \cdots + a_n(x_n - p_n) = 0$$

con $(a_1, a_2, ..., a_n) \neq (0, 0, ..., 0)$ (los a_i son los n menores de orden $n-1$ que se pueden formar con las $n-1$ primeras columnas; no pueden ser todas ceros ya que estas columnas son independientes por ser las columnas de coordenadas de vectores linealmente independientes).

Llamando $a_0 = -a_1p_1 - a_2p_2 - \cdots - a_np_n$, se obtiene la anterior ecuación [2] del enunciado.

Comprobemos finalmente que una ecuación como la [2] es ecuación de un hiperplano. Como los $a_1, a_2, ..., a_n$ no pueden ser todos nulos, podemos suponer que es $a_n \neq 0$. Recurriendo a los parámetros $\lambda_1 = x_1, \lambda_2 = x_2, ..., \lambda_{n-1} = x_{n-1}$, el lugar geométrico que representa la ecuación [2] puede expresarse mediante las siguientes ecuaciones paramétricas

$$\begin{cases} x_1 = 0 + \lambda_1 \\ x_2 = 0 + \lambda_2 \\ x_{n-1} = 0 + \lambda_{n-1} \\ x_n = -\dfrac{a_0}{a_n} - \dfrac{a_1}{a_n}\lambda_1 - \dfrac{a_2}{a_n}\lambda_2 - \cdots - \dfrac{a_{n-1}}{a_n}\lambda_{n-1} \end{cases}$$

Estas son las ecuaciones de una variedad afín que tiene por dirección al subespacio vectorial que engendran los vectores:

$$\bar{u}_1(1, 0, ..., 0, -a_1/a_n)$$
$$\bar{u}_2(0, 1, ..., 0, -a_1/a_n)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\bar{u}_{n-1}(0, 0, ..., 1, -a_{n-1}/a_n)$$

Como estos $n-1$ vectores son linealmente independientes, resulta que la ecuación [2] representa a una variedad lineal de dimensión $n-1$, es decir, a un hiperplano.

ECUACIONES CARTESIANAS DE UNA VARIEDAD LINEAL

[264]

Sea $P+U$ una variedad lineal, de un espacio afín euclídeo E (asociado al espacio vectorial V), y sea $U^{\perp}$ el subespacio ortogonal de U, que tiene a $(\bar{a}_1, ..., \bar{a}_k)$ como base. Se verifica que $P+U$ está formada por los puntos $X \in E$ que verifican a las siguientes ecuaciones:

$$\bar{a}_j \cdot \overrightarrow{PX} = 0 \quad \text{para } j = 1, ..., k$$

Si dim $E = n$ y recurriendo a las coordenadas en una referencia rectangular de E, a las que llamaremos $P(p_i)$, $X(x_i)$, $\bar{a}_j(a_{ji})$, entonces la variedad $P+U$ tiene por ecuaciones (cartesianas) a las:

$$a_{j1}(x_1 - p_1) + a_{j2}(x_2 - p_2) + \cdots + a_{jn}(x_n - p_n) = 0 \quad \text{para } j = 1, ..., k$$

donde las k filas $(a_{j1}, a_{j2}, ..., a_{jn})$ forman una matriz de rango k.

En particular, dados un punto $P(p_i)$ y un vector no nulo $\bar{a}(a_i)$, hay un y sólo un hiperplano que pasa por P y es ortogonal a $\bar{a}$, el cual está formado por los puntos $X(x_i)$ tales que:

$$\bar{a} \cdot \overrightarrow{PX} = 0 \quad \text{o} \quad a_1(x_1 - p_1) + a_2(x_2 - p_2) + \cdots + a_n(x_n - p_n) = 0$$

COMPROBACIONES

1. Un punto $X \in E$ es de la variedad $P+U$ si $\overrightarrow{PX} \in U$, esto es, si $\overrightarrow{PX}$ es ortogonal a $U^{\perp}$, que es tanto como decir $\bar{v} \cdot \overrightarrow{PX} = 0$ para todo $\bar{v} \in U^{\perp}$; como $(\bar{a}_1, ..., \bar{a}_k)$ es una base de $U^{\perp}$, la última relación se verifica si y sólo si $\bar{a}_j \cdot \overrightarrow{PX} = 0$ para $j = 1, ..., k$.
2. Al utilizar coordenadas cartesianas rectangulares, como la matriz métrica del producto escalar es, así, la unidad, resulta que la relación $\bar{a}_j \cdot \overrightarrow{PX} = 0$ toma la forma:

$$a_{j1}(x_1 - p_1) + a_{j2}(x_2 - p_2) + \cdots + a_{jn}(x_n - p_n) = 0$$

   Nótese que, como los vectores $\bar{a}_1, ..., \bar{a}_k$ son linealmente independientes, también lo son sus filas de coordenadas; esto es, la matriz $[a_{ji}]$, de tamaño $k \times n$, tiene rango k.
3. Si $P+U$ es un hiperplano, entonces $U^{\perp}$ es un subespacio unidimensional, engendrado por el vector $\bar{a}$; por ello, en este caso, las k ecuaciones anteriores se reducen a una sola, que es la $\bar{a} \cdot \overrightarrow{PX} = 0$.

OBSERVACION

Según se acaba de obtener, si la dimensión del espacio E es n, toda variedad lineal de dimensión h queda representada por un sistema de $k = n - h$ ecuaciones lineales del tipo $\sum_i a_{ji}x_i = b_j$ $(i = 1, ..., n; j = 1, ..., k)$, que matricialmente se pone $AX = B$, donde $A = [a_{ji}]$ es de tamaño $k \times n$ y tiene rango k. Si E no fuese euclídeo, se llega a la misma conclusión recurriendo al producto escalar que hace ortonormal a la base de la referencia que se tome en E.

β CUESTIONES DE CARACTER AFIN

Se analizan aquí algunos de los problemas de tipo afín, entre puntos y variedades lineales de un espacio puntual. Nos vamos a referir, pues, a las cuestiones relacionadas con la dependencia lineal, las intersecciones, la suma de variedades, el paralelismo. Aquí no tienen acomodo las cuestiones relacionadas con el producto escalar.

12.5. INTERSECCION Y SUMA DE VARIEDADES LINEALES

INTERSECCION DE VARIEDADES LINEALES

Dos o más variedades pueden tener intersección vacía; vamos a empezar caracterizando este hecho. Cuando su intersección no sea vacía, comprobaremos que ella es una variedad lineal, la cual determinaremos:

[265]

Sean $F_1 = P_1 + U_1$ y $F_2 = P_2 + U_2$ dos variedades lineales, del espacio afín E (asociado al espacio vectorial V). Se verifica que:

1. **$F_1 \cap F_2 \neq \emptyset$ si y sólo si $\overrightarrow{P_1P_2} \in U_1 + U_2$.**
2. **Si $P \in F_1 \cap F_2$, entonces $F_1 \cap F_2 = P + (U_1 \cap U_2)$.**

DEMOSTRACION

1. Si $P \in F_1 \cap F_2$, entonces existen $\vec{u}_1 \in U_1$ y $\vec{u}_2 \in U_2$ tales que $\overrightarrow{P_1P} = \vec{u}_1$ y $\overrightarrow{P_2P} = \vec{u}_2$ y, por tanto $\overrightarrow{P_1P_2} = \overrightarrow{PP_2} - \overrightarrow{PP_1} = \vec{u}_2 - \vec{u}_1$ es un vector de $U_1 + U_2$. Recíprocamente, supongamos ahora que $\overrightarrow{P_1P_2} \in U_1 + U_2$, lo que nos permite asegurar que existen $\vec{u}_1 \in U_1$ y $\vec{u}_2 \in U_2$ tales que $\overrightarrow{P_1P_2} = \vec{u}_1 - \vec{u}_2$; re-

sulta entonces que el punto $P = P_1 + \bar{u}_1 \in F_1$ se puede poner también en la forma:

$$P = P_1 + \bar{u}_1 = P_1 + (\overrightarrow{P_1P_2} + \bar{u}_2) = (P_1 + \overrightarrow{P_1P_2}) + \bar{u}_2 = P_2 + \bar{u}_2 \in F_2$$

luego $P \in F_1 \cap F_2$, por lo que $F_1 \cap F_2 \neq \emptyset$.

2. Recurriendo al punto P, se puede poner $F_1 = P + U_1$ y $F_2 = P + U_2$ y de ahí que se tenga:

$$X \in F_1 \cap F_2 \Leftrightarrow [\overrightarrow{PX} \in U_1 \text{ y } \overrightarrow{PX} \in U_2] \Leftrightarrow \overrightarrow{PX} \in U_1 \cap U_2 \Leftrightarrow$$
$$\Leftrightarrow X \in P + (U_1 \cap U_2)$$

OBSERVACION

La anterior propiedad [265], 2, se verifica también si, en lugar de dos, hubiera una familia cualquiera de variedades lineales; es decir, se verifica que: si P es un punto que pertenece a todas las variedades lineales $P_i + U_i$, para $i \in I$ (familia cualquiera), entonces

$$\bigcap_{i \in I} [P_i + U_i] = P + \bigcap_{i \in I} U_i$$

lo que se demuestra siguiendo los mismos pasos que antes.

EJERCICIO

En un espacio afín E, asociado al espacio vectorial V, se consideran dos variedades lineales $P_1 + U_1$ y $P_2 + U_2$. Compruébese que si sus direcciones U_1 y U_2 son subespacios suplementarios de V, entonces

$$(P_1 + U_1) \cap (P_2 + U_2)$$

consta de un y sólo un punto.

RESOLUCION

Como $U_1 + U_2 = V$, se verifica que $\overrightarrow{P_1P_2} \in U_1 + U_2$ y, de acuerdo con [265], 1, de ahí se deduce que $F_1 \cap F_2 \neq \emptyset$, luego existe un punto $P \in F_1 \cap F_2$. Por otra parte, como $U_1 \cap U_2 = O$, según [265], 2, ha de ser $F_1 \cap F_2 = P + O$, esto es $F_1 \cap F_2$ se reduce al punto P.

SUMA DE VARIEDADES LINEALES

La unión de dos variedades lineales no es, en general, una variedad lineal. Así, por ejemplo, la unión de un plano y una recta no incluida en él no es una

variedad afín. Resulta entonces que la unión de dos variedades lineales, que es el menor conjunto que incluye a ambas, no es en general la menor variedad lineal que incluye a aquellas. Pues bien, a esta variedad mínima se la llama suma de las variedades dadas:

[266]

Sean $F_1 = P_1 + U_1$ y $F_2 = P_2 + U_2$ dos variedades lineales, de un espacio afín E (asociado al espacio vectorial V). Se llama suma de F_1 y F_2 a la variedad afín, $F_1 + F_2$, engendrada por el conjunto $F_1 \cup F_2$, esto es, a la menor de las variedades lineales que incluyen a F_1 y a F_2.

La anterior suma $F_1 + F_2$, que pasa por los puntos P_1 y P_2, tiene por dirección al subespacio:

$$U = \mathcal{V}(\overrightarrow{P_1P_2}) + U_1 + U_2$$

(nótese que $U = U_1 + U_2$ si y sólo si $F_1 \cap F_2 \neq \emptyset$).

DEMOSTRACION

Como la variedad lineal $F = F_1 + F_2$ es la menor de cuantas contienen a F_1 y a F_2, su dirección U será, según lo dicho en [261], el subespacio vectorial engendrado por $C = \{\overrightarrow{XY} / X, Y \in F_1 \cup F_2\}$. Ahora bien: si $X, Y \in F_1$, entonces $\overrightarrow{XY}$ recorre todo U_1; si $X, Y \in F_2$, entonces $\overrightarrow{XY}$ recorre todo U_2; si $X \in F_1$ e $Y \in F_2$, entonces se puede poner $\overrightarrow{XY} = \overrightarrow{XP_1} + \overrightarrow{P_1P_2} + \overrightarrow{P_2Y}$, lo que permite asegurar que $\overrightarrow{XY}$ recorre todo el conjunto $\overrightarrow{P_1P_2} + U_1 + U_2$. Así pues, $C = \overrightarrow{P_1P_2} + U_1 + U_2$ y, por tanto, $U = \mathcal{V}(\overrightarrow{P_1P_2}) + U_1 + U_2$.

Nótese, finalmente, que la condición necesaria y suficiente para que sea $U = U_1 + U_2$ es que $\overrightarrow{P_1P_2} \in U_1 + U_2$, lo cual equivale, según se dijo en [265], 1, a que $F_1 \cap F_2 \neq \emptyset$.

OBSERVACION

La suma de más de dos variedades lineales se define, también, como la variedad lineal engendrada por la unión de todas ellas, o lo que es igual, como la menor variedad lineal que las incluye. Aplicando reiteradamente el resultado precedente, [266], se llega fácilmente a que la suma de varias variedades lineales es:

$$\sum_{i=0}^{n} (P_i + U_i) = P_0 + [\mathcal{V}(\overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_n}) + U_0 + U_1 + \cdots + U_n]$$

FORMULAS DE GRASSMANN (O DE LAS DIMENSIONES)

[267]

Si F_1 y F_2 son dos variedades lienales de dimensión finita, de un espacio afín E (asociado al espacio vectorial V), se verifica que:

1. $F_1 \cap F_2 \neq \emptyset \Rightarrow \dim F_1 + \dim F_2 = \dim (F_1 + F_2) + \dim (F_1 \cap F_2)$.
2. $F_1 \cap F_2 = \emptyset \Rightarrow \dim F_1 + \dim F_2 = \dim (F_1 + F_2) + \dim (F_1 \cap F_2) - 1$.

DEMOSTRACION

Llamemos $F_1 = P_1 + U_1$ y $F_2 = P_2 + U_2$. Como la dimensión de una variedad lineal es la dimensión de su dirección y de acuerdo con [266], será:

$$\dim (F_1 + F_2) = \dim (\mathcal{V}(\overrightarrow{P_1P_2}) + U_1 + U_2)$$

Recurriendo a la fórmula de Grassmann para subespacios vectoriales (véase [078]), se obtiene:

1. Si $F_1 \cap F_2 \neq \emptyset$, como entonces $\overrightarrow{P_1P_2} \in U_1 + U_2$, resulta que

$$\mathcal{V}(\overrightarrow{P_1P_2}) \subset U_1 + U_2, \text{ luego } \mathcal{V}(\overrightarrow{P_1P_2}) + U_1 + U_2 = U_1 + U_2, \text{ y}$$

$$\dim (F_1 + F_2) = \dim (U_1 + U_2) = \dim U_1 + \dim U_2 - \dim (U_1 \cap U_2) = \\ = \dim F_1 + \dim F_2 - \dim (F_1 \cap F_2)$$

2. Si $F_1 \cap F_2 = \emptyset$, entonces $\overrightarrow{P_1P_2} \notin U_1 + U_2$, lo que conduce a:

$$\dim (F_1 + F_2) = \dim (\mathcal{V}(\overrightarrow{P_1P_2}) + U_1 + U_2) = 1 + \dim (U_1 + U_2) = \\ = 1 + \dim U_1 + \dim U_2 - \dim (U_1 \cap U_2) = \\ = 1 + \dim F_1 + \dim F_2 - \dim (F_1 \cap F_2)$$

EJEMPLOS

La suma de dos rectas (distintas) es: un plano, si se cortan; un plano, si son paralelas; y una variedad tridimensional, si se cruzan. Esto se obtiene de aplicar la fórmula anterior, que (con la notación de antes) conduce, respectivamente, a:

$$\dim (F_1 + F_2) = \begin{cases} \dim F_1 + \dim F_2 - \dim (F_1 \cap F_2) = 1 + 1 - 0 = 2 \\ \dim F_1 + \dim F_2 - \dim (F_1 \cap F_2) + 1 = 1 + 1 - 1 + 1 = 2 \\ \dim F_1 + \dim F_2 - \dim (F_1 \cap F_2) + 1 = 1 + 1 - 0 + 1 = 3 \end{cases}$$

12.6. COORDENADAS BARICENTRICAS

BARICENTRO

[268]

Sean dados unos ciertos puntos P_1, P_2, ..., P_k, de un espacio afín E (asociado al espacio vectorial V); sean dados, también, unos escalares c_1, c_2, ..., c_k tales que $c_1 + c_2 + \cdots + c_k \neq 0$. Elegido arbitrariamente un punto O del espacio, sea P el punto definido por:

$$\overrightarrow{OP} = \frac{c_1\overrightarrow{OP_1} + c_2\overrightarrow{OP_2} + \cdots + c_k\overrightarrow{OP_k}}{c_1 + c_2 + \cdots + c_k} \qquad [1]$$

Este punto P no depende del origen O elegido y se llama *baricentro* de los puntos P_i afectados de los coeficientes c_i (para $i = 1, 2, ..., k$). Dicho baricentro queda definido por la relación:

$$c_1\overrightarrow{PP_1} + c_2\overrightarrow{PP_2} + \cdots + c_k\overrightarrow{PP_k} = \bar{o} \qquad [2]$$

DEMOSTRACION

Sea O' un nuevo punto origen, cualquiera. Recurriendo a este O', la relación [1] puede escribirse del siguiente modo:

$$\overrightarrow{OO'} + \overrightarrow{O'P} = \frac{\sum_{i=1}^{k} c_i(\overrightarrow{OO'} + \overrightarrow{O'P_i})}{\sum_{i=1}^{k} c_i} = \overrightarrow{OO'} + \frac{\sum_{i=1}^{k} c_i\overrightarrow{O'P_i}}{\sum_{i=1}^{k} c_i}$$

Restando, al primer y al tercer miembros, el vector $\overrightarrow{OO'}$, nos encontramos nuevamente con la relación [1] pero ahora con el origen O' en lugar del O; así pues, la definición de P no depende del origen empleado.

Para comprobar que [2] define al baricentro P, basta con tomar, en [1], como origen O al propio punto P.

EJERCICIO

Sea P el baricentro de los puntos P_1, ..., P_k afectados de coeficientes c_1, ..., c_k, siendo $c_1 + \cdots + c_k = c \neq 0$; sea Q el baricentro de los puntos Q_1, ..., Q_h afectados de los coeficientes d_1, ..., d_h, siendo $d_1 + \cdots + d_h = d \neq 0$. Si es R el baricentro de los puntos P y Q afectados de coeficientes c y d (se supone

$c+d\neq 0$) y si es R' el baricentro de los puntos $P_1, ..., P_k, Q_1, ..., Q_h$ afectados de los coeficientes $c_1, ..., c_k, d_1, ..., d_h$, pruébese que ha de ser $R = R'$.

RESOLUCION

De acuerdo con la definición de baricentro, [268], se puede poner (las siguientes sumas lo son para $i = 1, ..., k$ y $j = 1, ..., h$)

$$\overrightarrow{OR'} = \frac{\Sigma c_i \overrightarrow{OP_i} + \Sigma d_j \overrightarrow{OQ_j}}{\Sigma c_i + \Sigma d_j} = \frac{1}{c+d}\left(c\,\frac{\Sigma c_i \overrightarrow{OP_i}}{c} + d\,\frac{\Sigma d_j \overrightarrow{OQ_j}}{d}\right) = \frac{c\overrightarrow{OP} + d\overrightarrow{OQ}}{c+d} = \overrightarrow{OR}$$

de donde se infiere que $R' = R$.

DEPENDENCIA E INDEPENDENCIA LINEAL DE PUNTOS

Dos puntos son independientes si son distintos; tres puntos son independientes si no están alineados; cuatro puntos son independientes si no son coplanarios; k puntos son independientes si no están en una misma variedad lineal de dimensión $k-2$. Más detalladamente, se conviene en que:

[269]

En un espacio afín E, asociado al espacio vectorial V, se dice que los $k+1$ puntos $P_0, P_1, ..., P_k$ son *linealmente independientes* o *linealmente dependientes* según que los k vectores $\overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_k}$ sean, respectivamente, linealmente independientes o linealmente dependientes (esta definición no depende de cual sea el punto P_0, de ellos, que se tome como origen de los vectores).

Los puntos $P_0, P_1, ..., P_k$ son linealmente dependientes si, y sólo si, existen ciertos escalares $\alpha_0, \alpha_1, ..., \alpha_k$ tales que:

$$\begin{cases} \bullet\ \alpha_0\overrightarrow{OP_0} + \alpha_1\overrightarrow{OP_1} + \cdots + \alpha_k\overrightarrow{OP_k} = \bar{o}, \text{ para cualquier } O \in E \\ \bullet\ \alpha_0 + \alpha_1 + \cdots + \alpha_k = 0 \quad \text{y} \quad (\alpha_0, \alpha_1, ..., \alpha_k) \neq (0, 0, ..., 0) \end{cases} \qquad [1]$$

DEMOSTRACION

1. Empecemos comprobando que la definición anterior no depende del punto P_0 que se tome como origen de los vectores; para ello, será suficiente con probar que, si $\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, ..., \overrightarrow{P_0P_k}$ son linealmente dependientes, entonces también lo son $\overrightarrow{P_1P_0}, \overrightarrow{P_1P_2}, ..., \overrightarrow{P_1P_k}$. Supongamos, pues, que para ciertos escalares $\lambda_1, \lambda_2, ..., \lambda_k$ sea

$$\lambda_1\overrightarrow{P_0P_1} + \lambda_2\overrightarrow{P_0P_2} + \cdots + \lambda_k\overrightarrow{P_0P_k} = \bar{o} \quad \text{con} \quad (\lambda_1, \lambda_2, ..., \lambda_k) \neq (0, 0, ..., 0)$$

Poniendo $\overrightarrow{P_0P_i} = \overrightarrow{P_0P_1} + \overrightarrow{P_1P_i}$, la anterior relación toma la forma

$$\lambda_0\overrightarrow{P_1P_0} + \lambda_2\overrightarrow{P_1P_2} + \cdots + \lambda_k\overrightarrow{P_1P_k} = \bar{o}, \quad \text{donde} \quad \lambda_0 = -(\lambda_1 + \lambda_2 + \cdots + \lambda_k)$$

lo que prueba que $\overrightarrow{P_1P_0}$, $\overrightarrow{P_1P_2}$, ..., $\overrightarrow{P_1P_k}$ son linealmente dependientes, ya que los λ_0, λ_2, ..., λ_k no pueden ser todos nulos por no serlo los λ_1, λ_2, ..., λ_k.

2. Suponiendo ahora que P_0, P_1, ..., P_k son linealmente dependientes, comprobemos que se verifica la relación [1] para ciertos escalares α_0, α_1, ..., α_k. De acuerdo con la definición de dependencia lineal, existen unos escalares λ_1, λ_2, ..., λ_k no todos nulos y tales que

$$\lambda_1\overrightarrow{P_0P_1} + \lambda_2\overrightarrow{P_0P_2} + \cdots + \lambda_k\overrightarrow{P_0P_k} = \bar{o}$$

Para cualquiera que sea el punto $O \in E$, poniendo que $\overrightarrow{P_0P_i} = \overrightarrow{OP_i} - \overrightarrow{OP_0}$, la anterior relación queda en la forma

$$-(\lambda_1 + \lambda_2 + \cdots + \lambda_k)\overrightarrow{OP_0} + \lambda_1\overrightarrow{OP_2} + \cdots + \lambda_k\overrightarrow{OP_k} = \bar{o}$$

lo que prueba que se cumple [1] para

$$\alpha_0 = -(\lambda_1 + \lambda_2 + \cdots + \lambda_k) \quad . \quad \alpha_1 = \lambda_1, \ldots, \alpha_k = \lambda_k$$

que no son todos nulos y suman cero.

3. Recíprocamente, suponiendo que se cumple [1], comprobemos que los puntos P_0, P_1, ..., P_k son linealmente dependientes. Así es, ya que poniendo $\overrightarrow{OP_i} = \overrightarrow{P_0P_i} - \overrightarrow{P_0O}$, se obtiene que

$$\sum_{i=0}^{k} \alpha_i\overrightarrow{OP_i} = \bar{o} \quad \Rightarrow \quad \sum_{i=1}^{k} \alpha_i\overrightarrow{P_0P_i} = \left(\sum_{i=0}^{k} \alpha_i\right)\overrightarrow{P_0O} = 0\overrightarrow{P_0O} = \bar{o}$$

luego los puntos dados son linealmente dependientes, pues los α_1, α_2, ..., α_k no pueden ser todos nulos ya que no lo eran los α_0, α_1, α_2, ..., α_k y la suma de todos éstos es nula.

EJERCICIO

Sean dados los puntos P_1, P_2, ..., P_k, de un espacio afín E; sean α_1, α_2, ..., α_k unos escalares tales que $\alpha_1 + \alpha_2 + \cdots + \alpha_k = 0$. Pruébese que el vector

$$\bar{v} = \alpha_1\overrightarrow{OP_1} + \alpha_2\overrightarrow{OP_2} + \cdots + \alpha_k\overrightarrow{OP_k}$$

no depende del punto $O \in E$.

RESOLUCION

Recurriendo a otro punto $O' \in E$, se puede poner:

$$\bar{v} = \sum_{i=1}^{k} \alpha_i\overrightarrow{OP_i} = \sum_{i=1}^{k} \alpha_i(\overrightarrow{OO'} + \overrightarrow{O'P_i}) = \left(\sum_{i=1}^{k} \alpha_i\right)\overrightarrow{OO'} + \sum_{i=1}^{k} \alpha_i\overrightarrow{O'P_i} = \sum_{i=1}^{k} \alpha_i\overrightarrow{O'P_i}$$

EJERCICIO

Sea E un espacio afín de dimensión finita n, asociado al espacio vectorial V. Pruébese que:

1. En E el número máximo de puntos linealmente independientes es $n+1$.
2. Si $P_0, P_1, ..., P_k$ son puntos linealmente independientes de E, existen en E ciertos puntos $P_{k+1}, ..., P_n$ tales que unidos a los anteriores forman un conjunto de $n+1$ puntos independientes.

RESOLUCION

1. Si k puntos $P_1, P_2, ..., P_k$, de E, son linealmente independientes, entonces los $k-1$ vectores $\overrightarrow{P_1P_2}, ..., \overrightarrow{P_1P_k}$, de V, son linealmente independientes; como la dimensión de V es n, en V no puede haber más de n vectores linealmente independientes, luego $k-1 \leq n$, esto es, $k \leq n+1$.
2. Los vectores $\bar{u}_1 = \overrightarrow{P_0P_1}, ..., \bar{u}_k = \overrightarrow{P_0P_k}$ son linealmente independientes y, como V tiene dimensión n, existen unos vectores $\bar{u}_{k+1}, ..., \bar{u}_n$ tales que $(\bar{u}_1, ..., \bar{u}_k, \bar{u}_{k+1}, ..., \bar{u}_n)$ es base de V. Resulta evidente que los puntos $P_{k+1} = P_0 + \bar{u}_{k+1}, ..., P_n = P_0 + \bar{u}_n$ son solución del problema.

REFERENCIA AFIN. COORDENADAS BARICENTRICAS

[270]

Sea E un espacio afín de dimensión finita n, asociado al espacio vectorial V. Se llama *referencia afín* (o baricéntrica) de E a cualquier sucesión $(A_0, A_1, ..., A_n)$ de $n+1$ puntos linealmente independientes de E.

Dada una referencia afín $(A_0, A_1, ..., A_n)$, se verifica que para cada punto X de E, si se toma arbitrariamente un punto $O \in E$, existen unos únicos escalares $x_0, x_1, ..., x_n$ tales que

$$\begin{cases} \overrightarrow{OX} = x_0\overrightarrow{OA_0} + x_1\overrightarrow{OA_1} + \cdots + x_n\overrightarrow{OA_n} \\ x_0 + x_1 + \cdots + x_n = 1 \end{cases} \qquad [1]$$

Los $n+1$ escalares $x_0, x_1, ..., x_n$ no dependen del punto O que se elija (sólo dependen de la referencia afín y del punto X) y reciben el nombre de *coordenadas baricéntricas* del punto X en la referencia afín dada.

DEMOSTRACION

1. Veamos primero que existen unos escalares $x_0, x_1, ..., x_n$ que satisfacen a las relaciones [1]. Como la dimensión de E es n, todo sistema de $n+2$ puntos, como el formado por $A_0, A_1, ..., A_n, X$, es linealmente dependiente

y, por tanto (véase [269]) existen ciertos escalares $\alpha_0, \alpha_1, ..., \alpha_n, \beta$ no todos nulos y tales que

$$\alpha_0 + \alpha_1 + \cdots + \alpha_n + \beta = 0$$

y

$$\alpha_0\overrightarrow{OA_0} + \alpha_1\overrightarrow{OA_1} + \cdots + \alpha_n\overrightarrow{OA_n} + \beta\overrightarrow{OX} = \bar{o} \qquad [2]$$

El escalar β no puede ser nulo pues, si lo fuese, haciendo $\beta = 0$ en las anteriores relaciones [2] nos encontraríamos con que $A_0, A_1, ..., A_n$ deberían ser linealmente dependientes, lo que es falso pues forman una referencia afín. Resulta entonces que dividiendo por β en las relaciones [2], lo cual es lícito por ser $\beta \neq 0$, y llamando $x_i = \alpha_i/\beta$ (para $i = 0, 1, ..., n$), se obtienen las relaciones [1] del enunciado, que se verifican, pues, para dichos x_i.

2. Veamos ahora que los $x_0, x_1, ..., x_n$ son únicos. Si hubiera otros $x_0', x_1', ..., x_n'$, distintos de los anteriores, para los que también se verificasen las relaciones [1], restando las relaciones [1] expresadas para los x_i de las que se obtienen para los x_i' y llamando $\alpha_i = x_i - x_i'$ ($i = 0, 1, ..., n$), nos encontraríamos con que se habría de verificar que, para ciertos α_i no todos nulos:

$$\bar{o} = \alpha_0\overrightarrow{OA_0} + \alpha_1\overrightarrow{OA_1} + \cdots + \alpha_n\overrightarrow{OA_n} \quad ; \quad \alpha_0 + \alpha_1 + \cdots + \alpha_n = 0$$

y, en consecuencia con lo dicho en [269], los puntos $A_0, A_1, ..., A_n$ habrían de ser linealmente dependientes, lo que es falso.

3. Finalmente, comprobemos que los escalares $x_0, x_1, ..., x_n$ no dependen del punto $O \in E$ que se elija. Considérese otro punto $O' \in E$; recurriendo a O', la primera de las relaciones [1] puede ponerse en la forma

$$\overrightarrow{OO'} + \overrightarrow{O'X} = \sum_{i=0}^{n} x_i(\overrightarrow{OO'} + \overrightarrow{O'A_i}) = \left(\sum_{i=0}^{n} x_i\right)\overrightarrow{OO'} + \sum_{i=0}^{n} x_i\overrightarrow{O'A_i} =$$

$$= \overrightarrow{OO'} + \sum_{i=0}^{n} x_i\overrightarrow{O'A_i}$$

es decir, para los mismos x_i que aparecen en [1], se verifica que

$$\overrightarrow{O'X} = x_0\overrightarrow{O'A_0} + x_1\overrightarrow{O'A_1} + \cdots + x_n\overrightarrow{O'A_n}$$

Por tanto, al cambiar de origen no cambian los escalares $x_0, x_1, ..., x_n$.

OBSERVACIONES

1. Sea dada una referencia afín $(A_0, A_1, ..., A_n)$, del espacio afín E. Las coordenadas baricéntricas de un punto X son los escalares $x_0, x_1, ..., x_n$ (véase [270]) si, y sólo si, X es el baricentro de los puntos $A_0, A_1, ..., A_n$ afectados de los coeficientes $x_0, x_1, ... x_n$ (véase [268]). Como, dada la anterior referencia afín, todo punto del plano tiene coordenadas baricéntri-

cas, y éstas son únicas, resulta que todo punto del plano es baricentro de los puntos de la referencia afectados de unos determinados, y únicos, coeficientes.

2. Sea $(A_0, A_1, ..., A_n)$ una referencia afín, del espacio afín E, y sean $x_0, x_1, ..., x_n$ las coordenadas baricéntricas de un punto X en la citada referencia. De acuerdo con [270] y tomando $O = A_0$, será

$$\overrightarrow{A_0X} = x_0\overrightarrow{A_0A_0} + x_1\overrightarrow{A_0A_1} + \cdots + x_n\overrightarrow{A_0A_n} = x_1\overrightarrow{A_0A_1} + \cdots + x_n\overrightarrow{A_0A_n}$$

De aquí se infiere que $x_1, ..., x_n$ son las coordenadas cartesianas de X en la referencia cartesiana $(A_0; \overrightarrow{A_0A_1}, ..., \overrightarrow{A_0A_n})$.

De lo dicho se infiere también que, si $x_1, ..., x_n$ son las coordenadas cartesianas de X en la referencia $(A_0; \overrightarrow{A_0A_1}, ..., \overrightarrow{A_0A_n})$, entonces las coordenadas baricéntricas de X en la referencia afín $(A_0, A_1, ..., A_n)$ son los escalares $x_0, x_1, ..., x_n$, donde $x_0 = 1 - (x_1 + x_2 + \cdots + x_n)$.

EJEMPLOS

1. El *punto medio* de dos puntos (distintos) dados A_0 y A_1, de un espacio afín real E, es el punto de E cuyas coordenadas baricéntricas en la referencia afín (A_0, A_1) son $x_0 = 1/2, x_1 = 1/2$.
2. El *baricentro* (ordinario; con coeficientes iguales) de tres puntos A_0, A_1, A_2 (que forman triángulo; independientes) es el punto de E cuyas coordenadas baricéntricas en la referencia afín (A_0, A_1, A_2) son $x_0 = 1/3$, $x_1 = 1/3, x_2 = 1/3$.

γ CUESTIONES DE CARACTER EUCLIDEO

Para concluir nuestro análisis de los espacios afines, vamos a considerar aquellos asuntos de mayor notoriedad de cuantos pueden abordarse con el producto escalar. Hablaremos de ortogonalidad, de distancias (mínimas) entre variedades, de superficies esféricas.

12.7. ORTOGONALIDAD Y MINIMA DISTANCIA

PERPENDICULAR A UNA VARIEDAD POR UN PUNTO

Recordemos que, en el espacio geométrico ordinario, dada una recta (o un plano), por un punto exterior se puede trazar una y sólo una recta perpendicular a la recta dada (o al plano dado) y que la corte. Esto mismo va a ocurrir, en general, cuando el espacio tenga dimensión finita n cualquiera, al considerar, en lugar de la recta (o del plano) anterior, una variedad lineal cualquiera:

[271]

En un espacio afín euclídeo E de dimensión finita, asociado al espacio vectorial V, consideramos un punto P y una variedad lineal $F = Q + U$, cuya dirección es U; se supone que $P \notin F$. Se verifica que:

1. Existe una única recta R que pasa por P, es perpendicular a F y corta a F; a ella se la llama *recta perpendicular* a F trazada desde P. El punto P_0 de intersección de R y F se llama *proyección ortogonal* de P sobre F.
2. El vector $\overrightarrow{QP_0}$ es la proyección ortogonal de $\overrightarrow{QP}$ sobre U (véase [145]).
3. La distancia de P a un punto X de F, $d(P, X)$, alcanza su valor mínimo para $X = P_0$.

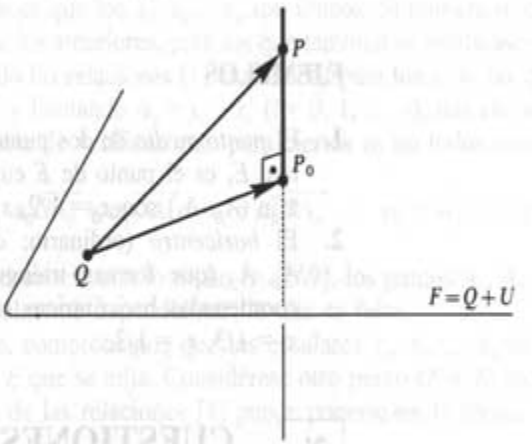


Figura 12.4.

DEMOSTRACION

Todo se basa en comprobar que existe un y sólo un punto P_0 de F tal que $\overrightarrow{PP_0}$ es ortogonal a U, lo que equivale a decir que el vector $\bar{x} = \overrightarrow{QP}$ se pueda expresar de manera única en la forma $\overrightarrow{QP} = \overrightarrow{QP_0} + \overrightarrow{P_0P}$ con $\overrightarrow{QP_0} \in U$ y $\overrightarrow{P_0P} \in U^{\perp}$. Ahora bien, esto último significa (véase [145]) que $\bar{x}_u = \overrightarrow{QP_0}$ es la proyección ortogonal de $\bar{x} = \overrightarrow{PQ}$ sobre U, la cual existe y es única (véase [146]) ya que V tiene dimensión finita.

Finalmente, para comprobar el último apartado, como el triángulo XP_0P es rectángulo en P_0, la longitud de su hipotenusa, $\|\overrightarrow{PX}\|$, es mayor que la del cateto $\|\overrightarrow{PP_0}\|$; es decir $d(P, X) > d(P, P_0)$ para cualquiera que sea el punto $X \neq P_0$ de F.

EJERCICIO

En un espacio afín euclídeo E de dimensión finita (asociado al espacio vectorial V), se consideran un punto P y una variedad lineal $F = Q + U$, donde U es la

dirección de F; se supone que P no pertenece a F. Se llama *variedad lineal perpendicular* a F trazada desde P a la variedad $F' = P + U^{\perp}$. Pruébese que F y F' se cortan en un solo punto y que este punto es la proyección ortogonal, P_0, de P sobre F.

RESOLUCION

Es evidente que la recta R perpendicular a F por P está incluida en F'; por tanto P_0, por pertenecer a R y a F, es un punto de la intersección de F y F'. Por tanto, según se dijo al estudiar la intersección de dos variedades lineales (véase [265], 2), será:

$$F \cap F' = P_0 + (U \cap U^{\perp}) = P_0 + O = \{P_0\}$$

DISTANCIA DE UN PUNTO A UNA VARIEDAD LINEAL

[272]

En un espacio afín euclídeo E de dimensión finita, asociado al espacio vectorial V, consideramos un punto P y una variedad lineal $F = Q + U$, cuya dirección es U; se supone que $P \notin F$. Se llama *distancia* de P a F a la menor de las distancias de P a los puntos de F y vale:

$$d(P, F) = \text{mín } \{d(P, X)/X \in F\} = \|\overrightarrow{PP_0}\|\}$$

(P_0 es la proyección ortogonal del punto P sobre la variedad F).

1. Si $(\bar{u}_1, \bar{u}_2, ..., \bar{u}_k)$ es una base ortogonal de U, entonces

$$d(P, F) = \left[d(P, Q)^2 - \sum_{i=1}^{k} \frac{(\overrightarrow{PQ} \cdot \bar{u}_i)^2}{\|\bar{u}_i\|^2} \right]^{1/2} \qquad [1]$$

2. Si $(\bar{v}_1, \bar{v}_2, ..., \bar{v}_h)$ es una base ortogonal de $U^{\perp}$, entonces

$$d = (P, F) = \left[\sum_{j=1}^{h} \frac{(\overrightarrow{PQ} \cdot \bar{v}_j)^2}{\|\bar{v}_j\|^2} \right]^{1/2} \qquad [2]$$

DEMOSTRACION

Comencemos observando que, según se dice en el apartado anterior, [271], la distancia $d(P, P_0) = \|\overrightarrow{PP_0}\|$ es la menor de todas las distancias $d(P, X)$ para $X \in F$ y, por ello, la definición que se ha dado es correcta:

1. Como el ángulo áng $(\overrightarrow{P_0P}, \overrightarrow{P_0Q})$ es recto, se verifica que

$$d(P, F) = \|\overrightarrow{PP_0}\| = (\|\overrightarrow{PQ}\|^2 - \|\overrightarrow{QP_0}\|^2)^{1/2} \qquad [3]$$

Por otra parte, como $\overrightarrow{QP_0} \in U$, este vector es combinación lineal de los $\bar{u}_i$, por lo que se puede poner, para ciertos escalares α_i:

$$\overrightarrow{QP_0} = \sum_{i=1}^{k} \alpha_i \bar{u}_i \quad \text{y} \quad \overrightarrow{PQ} = \overrightarrow{PP_0} - \sum_{i=1}^{k} \alpha_i \bar{u}_i$$

Multiplicando escalarmente por $\bar{u}_i$ los dos miembros de la última igualdad y teniendo en cuenta que $\overrightarrow{PP_0}$ es ortogonal a $\bar{u}_i$ y que la base $(\bar{u}_i)$ es ortogonal, se obtiene que

$$\overrightarrow{PQ} \cdot \bar{u}_i = -\alpha_i \|\bar{u}_i\|^2 \quad \text{para } i = 1, 2, \ldots, k$$

$$\|\overrightarrow{QP_0}\|^2 = \left\| \sum_{i=1}^{k} -\frac{\overrightarrow{PQ} \cdot \bar{u}_i}{\|\bar{u}_i\|^2} \bar{u}_i \right\|^2 = \sum_{i=1}^{k} \frac{(\overrightarrow{PQ} \cdot \bar{u}_i)^2}{\|\bar{u}_i\|^4} \|\bar{u}_i\|^2$$

Llevando esta última expresión a [3], se obtiene [1].

2. El vector $\overrightarrow{PP_0}$ es ortogonal a U y, por ello, existen ciertas escalares β_j tales que

$$\overrightarrow{PQ} = \overrightarrow{PP_0} + \overrightarrow{P_0Q} = (\beta_1 \bar{v}_1 + \cdots + \beta_h \bar{v}_h) + \overrightarrow{P_0Q}$$

Multiplicando escalarmente en esta igualdad por $\bar{v}_j$ (para $j = 1, 2, \ldots, h$), como $\overrightarrow{P_0Q}$ es ortogonal a $\bar{v}_j$ y como la base $(\bar{v}_j)$ es ortogonal, resulta que:

$$\overrightarrow{PQ} \cdot \bar{v}_j = \beta_j \|\bar{v}_j\|^2 \quad \text{para } j = 1, 2, \ldots, h$$

Por tanto, para la distancia $d(P, F)$ se obtiene:

$$d(P, F) = \|\overrightarrow{PP_0}\| = \left\| \sum_{j=1}^{h} \beta_j \bar{v}_j \right\| = \left[\sum_{j=1}^{h} \beta_j^2 \|\bar{v}_j\|^2 \right]^{1/2} = \left[\sum_{j=1}^{h} \frac{(\overrightarrow{PQ} \cdot \bar{v}_j)^2}{\|\bar{v}_j\|^2} \right]^{1/2}$$

EJERCICIO

En un espacio afín euclídeo E de dimensión n (asociado al espacio vectorial V), se consideran un punto P y el hiperplano H que pasa por el punto Q y es ortogonal al vector $\bar{a} \in V$. Compruébese que:

$$d(P, H) = \frac{|\bar{a} \cdot \overrightarrow{PQ}|}{\|\bar{a}\|}$$

Si en una referencia rectangular las coordenadas de P y la ecuación general de H son:

$$P(p_1, \ldots, p_n) \quad \text{y} \quad H: a_0 + a_1 x_1 + \cdots + a_n x_n = 0$$

obténgase que

$$d(P, H) = \frac{|a_0 + a_1 p_1 + \cdots + a_n p_n|}{\sqrt{a_1^2 + \cdots + a_n^2}}$$

RESOLUCION

Recurramos al resultado anterior [272], 2. En nuestro caso, $U^\perp = V(\bar{a})$ y la citada fórmula [2] queda en la forma:

$$d(P, H) = \left[\frac{(\overrightarrow{PQ} \cdot \bar{a})^2}{\|\bar{a}\|^2}\right]^{1/2} = \frac{|\overrightarrow{PQ} \cdot \bar{a}|}{\|\bar{a}\|}$$

Si la ecuación de H es la que se da en el enunciado, entonces $\bar{a}$ tiene por coordenadas a las $(a_1, ..., a_n)$. Por otra parte, para cualquiera que sea el punto $Q(q_i)$ de H, se puede poner:

$$\overrightarrow{PQ} \cdot \bar{a} = (q_1 - p_1)a_1 + \cdots + (q_n - p_n)a_n = (q_1a_1 + \cdots + q_na_n) - p_1a_1 - \cdots - p_na_n =$$
$$= -(a_0 + a_1p_1 + \cdots + a_np_n)$$

Por tanto, de la anterior expresión de $d(P, H)$ se obtiene que:

$$d(P, H) = \frac{|\overrightarrow{PQ} \cdot \bar{a}|}{\|\bar{a}\|} = \frac{|a_0 + a_1p_1 + \cdots + a_np_n|}{\sqrt{a_1^2 + \cdots + a_n^2}}$$

DISTANCIA ENTRE DOS VARIEDADES LINEALES

[273]

En un espacio afín euclídeo E de dimensión finita, asociado al espacio vectorial V, se consideran dos variedades lineales $F_1 = P_1 + U_1$ y $F_2 = P_2 + U_2$, cuyas direcciones son U_1 y U_2; sea $U = U_1 + U_2$.

1. Para cualesquiera que sean los puntos $X_1 \in F_1$ y $X_2 \in F_2$, la proyección ortogonal del vector $\overrightarrow{X_1X_2}$ sobre el subespacio $U^\perp$ es constante (no depende de los X_1 y X_2).
2. Existen dos puntos $Q_1 \in F_1$ y $Q_2 \in F_2$ tales que $\overrightarrow{Q_1Q_2}$ es la anterior proyección ortogonal.
3. Existe la menor de las distancias de un punto X_1 de F_1 a un punto X_2 de F_2; a ella se le llama distancia de F_1 a F_2 y vale:

$$d(F_1, F_2) = \text{mín } \{d(X_1, X_2)/X_1 \in F_1, X_2 \in F_2\} = \|\overrightarrow{Q_1Q_2}\|$$

DEMOSTRACION

1. Para cualesquiera que sean $X_1 \in F_1$ y $X_2 \in F_2$, pongamos:

$$\overrightarrow{X_1X_2} = \overrightarrow{X_1P_1} + \overrightarrow{P_1P_2} + \overrightarrow{P_2X_2} = \overrightarrow{P_1P_2} + (\overrightarrow{X_1P_1} + \overrightarrow{P_2X_2}) = \overrightarrow{P_1P_2} + \bar{u}$$

donde $\bar{u} \in U_1 + U_2$ ya que $\overrightarrow{X_1P_1} \in U_1$ y $\overrightarrow{P_2X_2} \in U_2$. Como la proyección ortogonal, sobre un subespacio, de una suma de vectores es igual a la suma de las proyecciones de los sumandos, resulta que

(Proy. sobre $U^\perp$ de $\overrightarrow{X_1X_2}$) = (Proy. sobre $U^\perp$ de $\overrightarrow{P_1P_2}$) + (Proy. sobre $U^\perp$ de $\bar{u}$)

y como el último sumando es nulo, pues $\bar{u} \in U$, resulta que la proyección de $\overrightarrow{X_1X_2}$ es, efectivamente, constante.

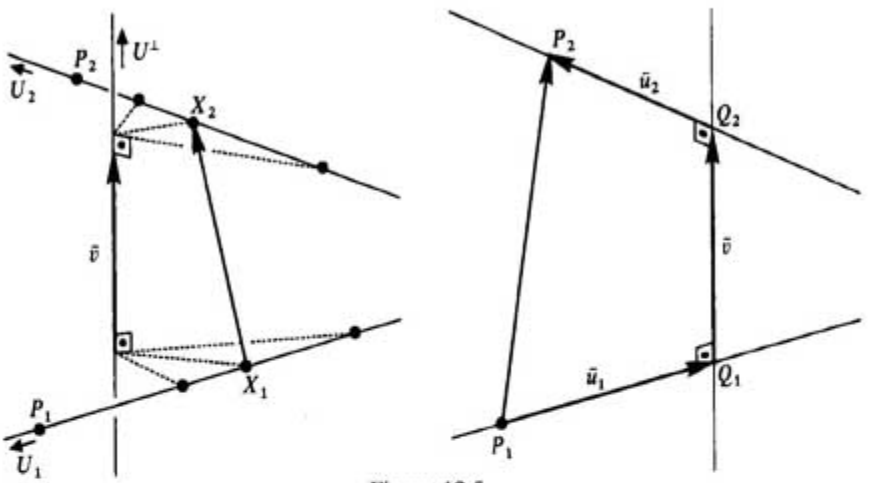


Figura 12.5.

2. Como los subespacios $U = U_1 + U_2$ y $U^\perp$ son suplementarios, el vector $\overrightarrow{P_1P_2}$ (como cualquier otro de V) puede expresarse de modo único en la forma $\overrightarrow{P_1P_2} = \bar{u} + \bar{v}$, donde $\bar{u} \in U$ y $\bar{v} \in U^\perp$; nótese que $\bar{v}$ es la proyección ortogonal de $\overrightarrow{P_1P_2}$ sobre $U^\perp$. Además, como $\bar{u} \in U_1 + U_2$, existen $\bar{u}_1 \in U_1$ y $\bar{u}_2 \in U_2$ tales que $\bar{u} = \bar{u}_1 + \bar{u}_2$. Así las cosas, los puntos $Q_1 = P_1 + \bar{u}_1$ y $Q_2 = P_2 - \bar{u}_2$ son los buscados, ya que, siendo obvio que $Q_1 \in F_1$ y $Q_2 \in F_2$, es $\overrightarrow{Q_1Q_2} = \bar{v}$, o sea, $\overrightarrow{Q_1Q_2}$ es la proyección ortogonal del apartado anterior.
3. La proyección ortogonal, $\bar{a}'$, de cualquier vector, $\bar{a}$, sobre un subespacio, es tal que $\|a'\| \leq \|\bar{a}\|$ y, por ello, para cualesquiera que sean $X_1 \in F_1$ y $X_2 \in F_2$, se verifica que $\|\overrightarrow{Q_1Q_2}\| \leq \|\overrightarrow{X_1X_2}\|$, donde $Q_1 \in F_1$ y $Q_2 \in F_2$ son los puntos hallados en el apartado anterior. Resulta, por ello, que

$$\text{mín } \{d(X_1, X_2)/X_1 \in F_1, X_2 \in F_2\} = \|\overrightarrow{Q_1Q_2}\|$$

OBSERVACION

Dadas dos variedades lineales $F_1 = P_1 + U_1$ y $F_2 = P_2 + U_2$, su distancia es, según acabamos de ver, la longitud de la proyección ortogonal del vector $\overrightarrow{P_1P_2}$ sobre el subespacio vectorial $U^\perp = (U_1 + U_2)^\perp$. Si $(\bar{e}_1, \bar{e}_2, ..., \bar{e}_s)$ es una base ortogonal de $U^\perp$, la proyección ortogonal de $\overrightarrow{P_1P_2}$ sobre $U^\perp$ es (véase [146]):

$$\sum_{i=1}^{s} \frac{\overrightarrow{P_1P_2} \cdot \bar{e}_i}{\|\bar{e}_i\|^2} \bar{e}_i$$

y, por tanto, la distancia entre F_1 y F_2 será:

$$d(F_1, F_2) = \left[\sum_{i=1}^{s} \frac{(\overrightarrow{P_1P_2} \cdot \bar{e}_i)^2}{\|\bar{e}_i\|^2} \right]^{1/2}$$

12.8. LA ESFERA

En un espacio afín euclídeo cualquiera, los puntos que están a una distancia dada de un punto fijo forman un lugar geométrico que se llama esfera. Estas variedades generalizan a las circunferencias, en el caso de dimensión dos, y a las esferas del espacio tridimensional ordinario.

DEFINICION

[274]

En un espacio afín euclídeo E, se llama *esfera* que tiene su centro en el punto $C \in E$ y cuyo radio es el número real $\rho > 0$ al conjunto formado por los puntos de E que distan ρ de C; esto es, a:

$$S = \{X \in E / \|\overrightarrow{CX}\| = \rho\}$$

Se llaman *diámetros* de una esfera a las rectas que pasan por su centro. Se dice que $X \in E$ es *interior* (*exterior*) a la esfera si $\|\overrightarrow{CX}\|$ es menor (mayor) que ρ.

Si E tiene dimensión finita n y en él se adopta una referencia cartesiana rectangular, la esfera de centro $C(c_1, c_2, ..., c_n)$ y radio $\rho > 0$ está formada por los puntos $X(x_1, x_2, ..., x_n)$ tales que:

$$(x_1 - c_1)^2 + (x_2 - c_2)^2 + \cdots + (x_n - c_n)^2 = \rho^2 \quad \text{(ecuación de la esfera)}$$

OBSERVACIONES

- Lo que aquí hemos llamado esfera, para otros es una «superficie esférica»; éstos llaman esfera a lo que para nosotros es el interior de una esfera, o conjunto de los puntos interiores. Este conjunto interior se suele llamar «bola abierta»; se llama «bola cerrada» al conjunto de los puntos situados en la esfera o dentro de la esfera.
- Un diámetro, de la esfera de centro C y radio ρ, es una recta de ecuación $X = C + \lambda\vec{u}$ (λ parámetro); si el vector $\vec{u}$ se toma unitario, el diámetro corta a la esfera en los puntos $C + \rho\vec{u}$ y $C - \rho\vec{u}$, que se llaman diametralmente opuestos.
- La ecuación de la esfera se obtiene de expresar que $\|\overrightarrow{CX}\|^2 = \rho^2$ utilizando coordenadas rectangulares; recurriendo a la expresión de la distancia (véase [256]), resulta trivial que la ecuación es:

$$(x_1 - c_1)^2 + (x_2 - c_2)^2 + \cdots + (x_n - c_n)^2 = \rho^2$$

operando en esta ecuación, se obtiene también:

$$(x_1^2 + \cdots + x_n^2) - 2(c_1x_1 + \cdots + c_nx_n) + (c_1^2 + \cdots + c_n^2 - \rho^2) = 0$$

INTERSECCION DE ESFERA E HIPERPLANO; TANGENCIA

[275]

En un espacio afín euclídeo E, se consideran la esfera S, que tiene centro C y radio $\rho > 0$, y un hiperplano H, que dista d del centro de S; sea C' la proyección ortogonal de C sobre H. La intersección de S y H es:

- Si $d < \rho$, entonces $S \cap H$ es la esfera de H que tiene su centro en C' y cuyo radio es $\rho' = \sqrt{\rho^2 - d^2}$; se dice que H es secante con S.
- Si $d > \rho$, entonces $S \cap H = \emptyset$; se dice que H es exterior a S.
- Si $d = \rho$, entonces $S \cap H$ se reduce al punto C'. En este caso, se dice que H es tangente, en C', a la esfera S.

Si P es un punto de la esfera S, el hiperplano que pasa por P y es ortogonal al vector $\overrightarrow{CP}$ es el *hiperplano tangente* en P a la esfera.

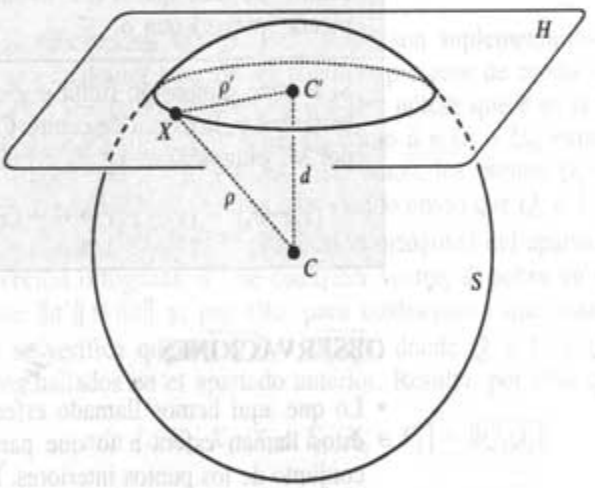


Figura 12.6.

DEMOSTRACION

El hiperplano H está formado por los puntos $X \in E$ tales que $\overrightarrow{C'X}$ es ortogonal a $\overrightarrow{CC'}$, esto es, que $\overrightarrow{C'X} \cdot \overrightarrow{CC'} = 0$. Los puntos de la esfera S son aquellos $X \in E$ tales que $\|\overrightarrow{CX}\| = \rho$, esto es, que $\overrightarrow{CX} \cdot \overrightarrow{CX} = \rho^2$. Por tanto $S \cap H$ está formada por aquellos puntos de H tales que

$$(\overrightarrow{CC'} + \overrightarrow{C'X}) \cdot (\overrightarrow{CC'} + \overrightarrow{C'X}) = \rho^2 \quad \text{o sea} \quad d^2 + \|\overrightarrow{C'X}\|^2 = \rho^2$$

Así, pues, $S \cap H$ la forman los puntos de H tales que $\|\overrightarrow{C'X}\| = \sqrt{\rho^2 - d^2}$, que es la esfera de H que tiene centro en C' y radio ρ' (si $d < \rho$). Si $d = \rho$, la anterior

relación toma la forma $\|\overrightarrow{C'X}\| = 0$, luego $S \cap H = \{C'\}$. Si $d > \rho$ no hay solución (real) de $\|\overrightarrow{C'X}\|^2 = \rho^2 - d^2$, con lo que $S \cap H = \emptyset$.

Finalmente, dado $P \in S$, el plano que pasa por P y es perpendicular a $\overrightarrow{CP}$ está a distancia $d = \rho$ de C, luego es tangente a S.

EJERCICIO

En un espacio afín euclídeo E, se considera la esfera S de centro en el punto C y radio ρ. Sea F una variedad lineal de dirección U que dista d del centro C, siendo $d < \rho$; se conoce la proyección ortogonal C' de C sobre F. Pruébese que la intersección $S \cap F$ es una esfera de F, hallando su centro y su radio.

RESOLUCION

Las variedades F y S están formadas por los puntos $X \in E$ tales que

$$F: \overrightarrow{C'X} \subset U \quad \text{y} \quad S: \|\overrightarrow{CX}\|^2 = \rho^2$$

Por tanto, $S \cap F$ está formado por aquellos puntos X de F tales que

$$(\overrightarrow{CC'} + \overrightarrow{C'X}) \cdot (\overrightarrow{CC'} + \overrightarrow{C'X}) = \rho^2$$

Ahora bien, como $\|\overrightarrow{CC'}\| = d$ y $\overrightarrow{CC'} \cdot \overrightarrow{C'X} = 0$, ya que $\overrightarrow{C'X} \subset U$ y $\overrightarrow{CC'} \perp U^{\perp}$, la última relación equivale a $\|\overrightarrow{C'X}\|^2 = \rho^2 - d^2$. Así pues, $S \cap F$ es la esfera de F que tiene centro en C' y cuyo radio es $\rho' = \sqrt{\rho^2 - d^2}$.

❑ INTERSECCION DE DOS ESFERAS

[276]

En un espacio afín euclídeo E, se consideran las esferas S_1 y S_2, que tienen sus centros en C_1 y C_2 y cuyos radios son ρ_1 y ρ_2, respectivamente; sea $d = \|\overrightarrow{C_1C_2}\|$. La intersección de S_1 y S_2 es:

1. Si con vértices en C_1 y C_2 se puede construir un triángulo cuyos lados sean d, ρ_1 y ρ_2, esto es, si $|\rho_1 - \rho_2| < d < (\rho_1 + \rho_2)$, llamando ρ a la altura de dicho triángulo correspondiente al lado $\overrightarrow{C_1C_2}$ y si C es el pie de dicha altura, entonces $S_1 \cap S_2$ es la esfera, del hiperplano ortogonal a $\overrightarrow{C_1C_2}$ por C, que tiene su centro en C y cuyo radio es ρ.
2. Si $\rho_1 + \rho_2 = d$ o si $|\rho_1 - \rho_2| = d$, entonces $S_1 \cap S_2$ se reduce a un punto.
3. En los demás casos, o sea, si $d > \rho_1 + \rho_2$ o si $d < |\rho_1 - \rho_2|$, la intersección es vacía, $S_1 \cap S_2 = \emptyset$.

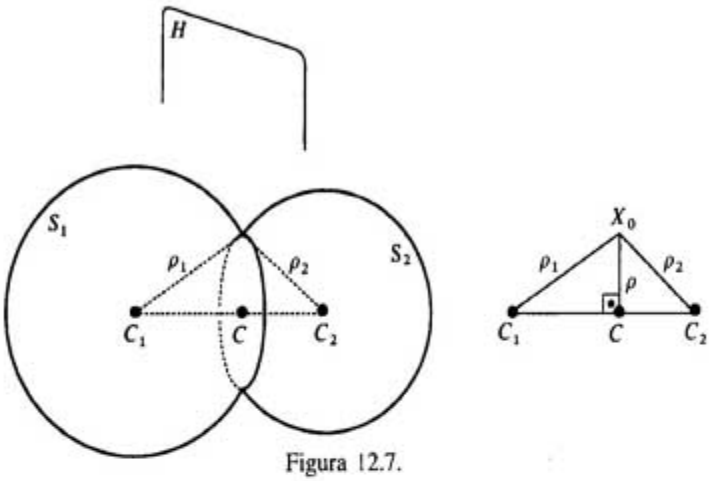


Figura 12.7.

DEMOSTRACION

1. Llamemos H al hiperplano que pasa por C y es ortogonal a $\overrightarrow{C_1C_2}$. Es de notar que tanto $S_1 \cap H$ como $S_2 \cap H$ son, ambos, la esfera de H que tiene centro en C y radio ρ. Por ello, hemos de probar que $S_1 \cap S_2$ es igual a $S_1 \cap H$ o a $S_2 \cap H$. Si tenemos en cuenta que S_1, S_2 y H son los lugares geométricos de los puntos $X \in E$ tales que:

$$S_1 : \|\overrightarrow{C_1X}\| = \rho_1, \quad S_2 : \|\overrightarrow{C_2X}\| = \rho_2, \quad H : \overrightarrow{CX} \cdot \overrightarrow{C_1C_2} = 0$$

tendremos que comprobar la siguiente equivalencia:

$$[\|\overrightarrow{C_1X}\| = \rho_1 \quad \text{y} \quad \|\overrightarrow{C_2X}\| = \rho_2] \Leftrightarrow [\overrightarrow{CX} \cdot \overrightarrow{C_1C_2} = 0 \quad \text{y} \quad \|\overrightarrow{C_2X}\| = \rho_2]$$

Para ello, recurriendo a que $\overrightarrow{C_1X} = \overrightarrow{C_1C} + \overrightarrow{CX}$ y a que $\overrightarrow{C_2X} = \overrightarrow{C_2C} + \overrightarrow{CX}$, notemos que

$$\begin{aligned} \|\overrightarrow{C_1X}\| = \rho_1 &\Leftrightarrow \|\overrightarrow{C_1C}\|^2 + 2\overrightarrow{CX} \cdot \overrightarrow{C_1C} + \|\overrightarrow{CX}\|^2 = \rho_1^2 \Leftrightarrow \\ &\Leftrightarrow 2\overrightarrow{CX} \cdot \overrightarrow{C_1C} + \|\overrightarrow{CX}\|^2 = \rho^2 \end{aligned} \qquad [1]$$

$$\begin{aligned} \|\overrightarrow{C_2X}\| = \rho_2 &\Leftrightarrow \|\overrightarrow{C_2C}\|^2 + 2\overrightarrow{CX} \cdot \overrightarrow{C_2C} + \|\overrightarrow{CX}\|^2 = \rho_2^2 \Leftrightarrow \\ &\Leftrightarrow 2\overrightarrow{CX} \cdot \overrightarrow{C_2C} + \|\overrightarrow{CX}\|^2 = \rho^2 \end{aligned} \qquad [2]$$

Por tanto, si es $\|\overrightarrow{C_1X}\| = \rho_1$ y $\|\overrightarrow{C_2X}\| = \rho_2$, restanto [2] de [1], se obtiene que $\overrightarrow{CX} \cdot \overrightarrow{C_1C_2} = 0$. Recíprocamente, si es $\|\overrightarrow{C_2X}\| = \rho_2$ y $\overrightarrow{CX} \cdot \overrightarrow{C_1C_2} = 0$, de [2] se deduce que $\|\overrightarrow{CX}\|^2 = \rho^2$, lo que pone de manifiesto que también se cumple [1], es decir que $\|\overrightarrow{C_1X}\| = \rho_1$.

2. Si $\rho_1 + \rho_2 = d$ o si $|\rho_1 - \rho_2| = d$, entonces también es de aplicación lo dicho en el punto anterior, pero ahora $\rho = 0$, con lo que la intersección se reduce a un punto (el punto C).

3. En los demás casos no hay ningún punto X en S_1 y en S_2, pues de haberlo, existiría el triángulo que se cita en el apartado 1, lo que no es posible tanto si es $d > \rho_1 + \rho_2$ como si es $d < |\rho_1 - \rho_2|$.

EJERCICIO

Hallar el centro C y el radio ρ de la intersección de las esferas S_1, de centro C_1 y radio ρ_1, y S_2, de centro C_2 y radio ρ_2. Llámese $d = \|\overrightarrow{C_1C_2}\|$; se supone que $|\rho_1 - \rho_2| < d < \rho_1 + \rho_2$.

RESOLUCION

En el triángulo $C_1C_2X_0$ de la Figura 11.7, aplicando el teorema del coseno correspondiente al vértice C_1, se obtiene que

$$\rho_2^2 = \rho_1^2 + d^2 - 2\rho_1 d \cos C_1$$

Recurriendo al $\cos C_1$, se obtiene que

$$C = C_1 + \overrightarrow{C_1C} = C_1 + \frac{\|\overrightarrow{C_1C}\|}{\|\overrightarrow{C_1C_2}\|}\overrightarrow{C_1C_2} = C_1 + \frac{\rho_1 \cos C_1}{d}\overrightarrow{C_1C_2} =$$

$$= C_1 + \frac{d^2 + \rho_1^2 - \rho_2^2}{2d^2}\overrightarrow{C_1C_2}$$

$$\rho^2 = \rho_1^2 - \|\overrightarrow{C_1C}\|^2 = \rho_1^2(1 - \cos^2 C_1) = \frac{1}{2}(\rho_1^2 + \rho_2^2) - \frac{1}{4d^2}(\rho_1^2 - \rho_2^2)^2 - \frac{d^2}{4}$$

INTERSECCION DE UNA ESFERA Y UNA RECTA; TANGENCIA

[277]

1. En un espacio afín euclídeo E, se consideran la esfera S, de centro en el punto C y radio ρ, y una recta r que dista d del centro C. La intersección $S \cap R$ consta de:

   - dos puntos si $d < \rho$; entonces se dice que r es secante con S.
   - un punto si $d = \rho$; entonces se dice que r es tangente a S.
   - ningún punto si $d > \rho$; entonces se dice que r es exterior a S.

2. Si P es un punto exterior a la esfera S y llamando $\delta = \|\overrightarrow{PC}\|$, una recta que pasa por P es tangente a S si, y sólo si, el ángulo θ que forma con la recta PC es tal que $\cos\theta = \sqrt{1 - (\rho/\delta)^2}$. Estas rectas engendran una variedad que se llama *hipercono tangente* a S desde P; ella está formada por los puntos $X \in E$ tales que

$$|\overrightarrow{PC} \cdot \overrightarrow{PX}| = \|\overrightarrow{PX}\|\sqrt{\delta^2 - \rho^2}$$

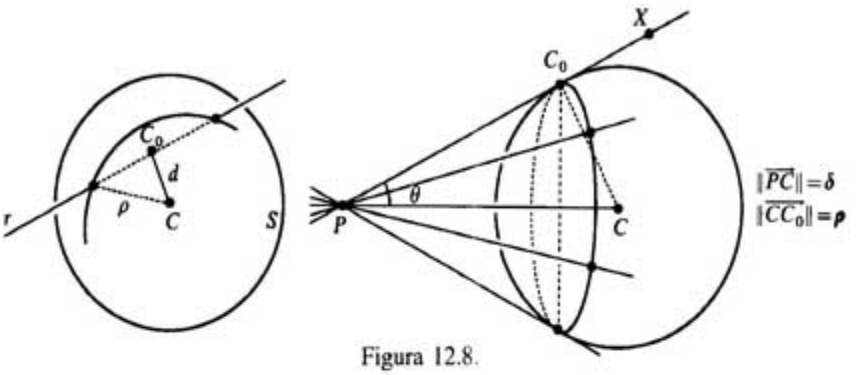


Figura 12.8.

DEMOSTRACION

1. Llamemos C_0 a la proyección ortogonal de C sobre r; se verifica que $\|\overrightarrow{CC_0}\| = d$. La esfera S y la recta r están formadas por los puntos $X \in \boldsymbol{E}$ tales que:

$$S: \|\overrightarrow{CX}\| = \rho \quad \text{y} \quad r: \overrightarrow{C_0X} = \lambda \bar{u} \quad (\lambda \in \mathbb{R})$$

donde $\bar{u}$ es un cierto vector unitario ortogonal a $\overrightarrow{CC_0}$. La intersección $S \cap r$ estará formada por aquellos puntos de r que se obtienen al tomar para $\boldsymbol{\lambda}$ las raíces de la ecuación

$$\rho^2 = \|\overrightarrow{CX}\|^2 = \|\overrightarrow{CC_0} + \overrightarrow{C_0X}\|^2 = \|\overrightarrow{CC_0} + \lambda\bar{u}\|^2 =$$
$$= \|\overrightarrow{CC_0}\|^2 + 2\lambda\bar{u} \cdot \overrightarrow{CC_0} + \lambda^2\|\bar{u}\| = d^2 + \lambda^2$$

Si es $d < \rho$, hay dos raíces, $\lambda = \pm\sqrt{\rho^2 - d^2}$, luego la intersección consta de dos puntos. Si $d = \rho$, hay una raíz, $\lambda = 0$, luego la intersección se reduce a un punto, que es el C_0. Si $d > \rho$, no hay soluciones reales y la intersección es vacía.

2. Una recta r, que pasa por P, será tangente, según se acaba de obtener, si la distancia de C a r es ρ, o sea, si $\|\overrightarrow{CC_0}\| = \rho$, o lo que es igual si el ángulo $\theta = \text{áng}(\overrightarrow{PC}, \overrightarrow{PC_0})$, es tal que

$$\cos\theta = \frac{\|\overrightarrow{PC_0}\|}{\|\overrightarrow{PC}\|} = \frac{\sqrt{\delta^2 - \rho^2}}{\delta} = \sqrt{1 - (\rho/\delta)^2}$$

Un punto X será de una de las rectas tangentes desde P a S si, y sólo si, el ángulo $\theta = \text{áng}(\overrightarrow{PC}, \overrightarrow{PX})$ tiene por coseno a $\pm\sqrt{1 - (\rho/\delta)^2}$, esto es:

$$\overrightarrow{PC} \cdot \overrightarrow{PX} = \|\overrightarrow{PC}\|\,\|\overrightarrow{PX}\| \cos\theta = \pm\delta\sqrt{1 - (\rho/\delta)^2}\,\|\overrightarrow{PX}\|$$

EJERCICIO

En un espacio afín euclídeo E, se considera la esfera S que tiene su centro en el punto C y cuyo radio es ρ. Considérese también la recta r que pasa por el punto P y tiene la dirección del vector $\bar{u}$ (unitario); sea $\delta = \|\overrightarrow{PC}\|$ y sea $\theta = \text{áng}\,(\overrightarrow{PC}, \bar{u})$. Hallar los puntos que forman la intersección $S \cap r$.

RESOLUCION

Los puntos que forman S y r son los $X \in E$ tales que:

$$S: \|\overrightarrow{CX}\| = \rho \quad \text{y} \quad r: \overrightarrow{PX} = \lambda \bar{u} \quad (\lambda \in \mathbb{R})$$

La intersección está formada por los puntos de r que se obtienen de aquellos $\lambda \in \mathbb{R}$ que verifican a:

$$\rho^2 = \|\overrightarrow{CX}\|^2 = \|\overrightarrow{CP} + \lambda\bar{u}\|^2 = \|\overrightarrow{CP}\|^2 + 2\lambda\bar{u} \cdot \overrightarrow{CP} + \lambda^2\|\bar{u}\|^2 =$$
$$= \delta^2 + 2\lambda\delta \cos\theta + \lambda^2$$

Las raíces de esta ecuación son las

$$\lambda = -\delta \cos\theta \pm \sqrt{\rho^2 - \delta^2 \operatorname{sen}^2\theta}$$

Por tanto, si $|\operatorname{sen}\theta| < \rho/\delta$ entonces hay dos puntos de intersección que son los

$$X = P + \left(-\delta \cos\theta \pm \sqrt{\rho^2 - \delta^2 \operatorname{sen}^2\theta}\right)\bar{u}$$

Si $|\operatorname{sen}\theta| = \rho/\delta$, entonces hay un punto de intersección, que es el $X = P - (\delta \cos\theta)\bar{u}$. Si $|\operatorname{sen}\theta| > \rho/\delta$, entonces la intersección es vacía.

Ejercicios y problemas a la parte V

ENUNCIADOS

V.1. En el plano afín y usando coordenadas cartesianas xy, se consideran dos puntos $P_1(\alpha_1, \beta_1)$ y $P_2(\alpha_2, \beta_2)$ y una recta $r: ax + by + c = 0$. Hallar la relación entre anteriores datos para que P_1 y P_2 estén a «distinto lado» de r (el segmento P_1P_2 corta a r).

V.2. En el plano afín, se consideran dos triángulos ABC y $A'B'C'$ tales que las tres rectas AA', BB' y CC' pasan por un mismo punto O. Sean P, Q y R los puntos de intersección de los siguientes pares de rectas: AB y $A'B'$, BC y $B'C'$ y CA y $C'A'$. Compruébese que los puntos P, Q y R están alineados (teorema de Desargues).

V.3. Dados tres puntos alineados P, Q y R (de E_1, de E_2 o de E_3), se llama «razón simple» (PQR) al número real ρ que permite poner $\overrightarrow{PR} = \rho\overrightarrow{PQ}$.

Sean A_1, A_2 y A_3 tres puntos no alineados del plano afín E_2 y sea r una recta que no pasa por ninguno de dichos puntos. Se llaman B_1, B_2 y B_3 a los puntos de intersección de r con las rectas A_2A_3, A_3A_1 y A_1A_2, respectivamente. Pruébese que (teorema de Menelao):

$$(B_1A_2A_3)(B_2A_3A_1)(B_3A_1A_2) = 1$$

V.4. Sea $A_1A_2A_3$ un triángulo del plano afín E_2 y sean B_1, B_2 y B_3 puntos situados en los segmentos A_2A_3, A_3A_1 y A_1A_2, respectivamente. Pruébese que las tres rectas A_1B_1, A_2B_2 y A_3B_3 pasan, las tres, por un mismo punto si y sólo si el producto de las razones simples (véase el problema precedente) $(B_1A_2A_3)$, $(B_2A_3A_1)$ y $(B_3A_1A_2)$ vale -1 (teorema de Ceva).

V.5. Se llama «cuadrilátero completo», del plano afín E_2, a la figura que forman cuatro rectas (lados) que se cortan dos a dos y tales que tres de ellas no pasan por un mismo punto; los seis puntos de intersección de cada dos lados se llaman vértices; las tres rectas que unen vértices opuestos (no situados en un mismo lado) se llaman diagonales. Pruébese que los puntos medios de las tres diagonales de un cuadrilátero completo están alineados.

V.6. En el plano afín E_2 y utilizando coordenadas cartesianas, se dan los puntos $A(-7, 3)$, $B(7, -4)$ y $G(5, 2)$. Hallar el punto $C \in E_2$ tal que el triángulo ABC tiene su baricentro en G.

V.7. Sean r y s dos rectas que se cortan en un punto O; sea A un punto coplanario con las rectas y no situado en ninguna de ellas; sea B el punto medio de OA. En r se toma un punto variable, P; sea Q la intersección de s con PB; sea R la intersección de AP con la paralela a OA por Q. Hallar el lugar geométrico que describe R (compruébese que la recta OA forma parte del lugar).

V.8. Compruébese que en un triángulo ABC cualquiera, las alturas pasan, las tres, por un mismo punto.

V.9. En el plano euclídeo y usando coordenadas rectangulares xy, se consideran los puntos $A(0, 0)$, $B(0,4)$ y $C(4, 2)$. Hallar el ortocentro, el baricentro y el circuncentro del triángulo ABC.

V.10. En el plano euclídeo y usando coordenadas rectangulares xy, se considera la recta $r: 3x + 4y + 7 = 0$ y el punto $P(-2, 1)$; sea C el cuadrado de centro en P, cuyos lados son paralelos y perpendiculares a r y cuyos lados miden 2. Hallar las ecuaciones de dichos lados.

V.11. En el plano euclídeo y usando coordenadas rectangulares Oxy, se consideran las rectas $y = 2x$ e $y = -2x$; sea A un punto variable de la primera y B uno de la segunda. Hallar el lugar geométrico que describe el baricentro del triángulo OAB si A y B se mueven de manera que el área del triángulo OAB valga 9.

V.12. En el plano euclídeo E_2 se considera una referencia cartesiana $(O; \bar{e}_1, \bar{e}_2)$ cuyos vectores $\bar{e}_1$ y $\bar{e}_2$ son unitarios y tales que áng $(\bar{e}_1, \bar{e}_2) = \theta$. Se considera otra referencia $(C; \bar{u}_1, \bar{u}_2)$ en la que $C(a, b)$, $\bar{u}_1$ y $\bar{u}_2$ son unitarios y tales que áng $(\bar{e}_1, \bar{u}_1) = \alpha$ y áng $(\bar{e}_1, \bar{u}_2) = \beta$. Relacionar (en función de a, b, θ, α y β) las coordenadas (x, y) de un punto en la primera referencia con las coordenadas (x', y') del mismo punto en la nueva referencia.

V.13. En el plano euclídeo E_2 y usando coordenadas rectangulares (x, y), se consideran las rectas que ad-

miten ecuación de la forma $x + 2\alpha y + \alpha^2 = 0$, donde $\alpha \in \mathbb{R}$ es un parámetro. Se pide:

1. Relación entre a y b para que por $P(a, b)$ pasen dos rectas de la familia.
2. Lugar geométrico de los puntos P por los que pase una sola recta.
3. Lugar geométrico de los puntos por los que pasan dos rectas perpendiculares.

V.14. Sea ABC un triángulo isósceles ($AB = AC$). Sean P y Q dos puntos variables que recorren AB y BC, respectivamente, de modo que la proyección ortogonal de PQ sobre BC tiene longitud mitad que BC. Pruébese que la perpendicular a PQ por Q pasa por un punto fijo.

V.15. En el plano euclídeo E_2 y usando coordenadas rectangulares xy, se considera el paralelogramo determinado por las cuatro rectas $ax + by \pm c = 0$ y $\alpha x + \beta y \pm \gamma = 0$. Hallar el área de dicho paralelogramo.

V.16. En el plano euclídeo E_2, se consideran dos rectas r y s que se cortan. Hallar el lugar geométrico de los puntos de E_2 tales que la suma de sus distancias a r y a s es constante y vale k.

V.17. En el plano euclídeo E_2 y usando coordenadas rectangulares, se consideran las circunferencias C y C' que tienen centros en los puntos $O(\alpha, \beta)$ y O' $(\alpha + d \cos \varphi, \beta + d \operatorname{sen} \varphi)$ y radios r y r'. Se pide:

1. Hallar una ecuación cuyas raíces sean las pendientes de las tangentes a C desde un punto $P(a, b)$.
2. Hallar las ecuaciones de las cuatro tangentes comunes a las circunferencias C y C' (se supone $r < r'$ y $r + r' < d$).

V.18. En el espacio afín E_3 se considera una cierta referencia cartesiana, en la que las coordenadas de un punto X son (x, y, z). Se considera otra referencia: la que tiene origen en $C(-2, 1, -1)$ y por vectores a $\vec{u}_1 = (1, 2, 3)$, $\vec{u}_2 = (0, 1, 1)$ y $\vec{u}_3 = (2, -3, 0)$. Se pide:

1. Hallar las coordenadas (x', y', z') de X en la nueva referencia.
2. Hallar los puntos $X \in E_3$ que tengan iguales coordenadas en ambas referencias.

V.19. En el espacio afín E_3 y usando coordenadas cartesianas xyz, se consideran un punto $P(-1, 1, 2)$, un plano $\pi: 3x + y - z = 9$ y dos rectas r y s:

$$r\begin{cases} x + y - z + 2 = 0 \\ 2x - 2y + z + 1 = 0 \end{cases} \qquad s\begin{cases} x + y - 2z - 2 = 0 \\ x - y + z + 4 = 0 \end{cases}$$

Se pide: 1) la recta que pasa por P y se apoya en (corta a) las rectas r y s; y 2) la recta que pasa por P se apoya en r y es paralela a π.

V.20. En el espacio afín E_3 y usando coordenadas cartesianas xyz, se consideran los cuatro planos siguientes:

$$x + 3y + z + 4 = 0 \qquad x + 6y + 2z + 8 = 0$$
$$x + 4y + 6 = 0 \qquad x + 8y + 2z + 10 = 0$$

Hallar los vértices del tetraedro cuyas caras son los planos dados.

V.21. Estudiar, en función de los valores del parámetro $\alpha \in \mathbb{R}$, la posición relativa de las rectas:

$$\frac{2x}{\alpha + 3} = \frac{y + 2}{1} = \frac{z - 2}{-1} \quad \text{y} \quad \begin{cases} 2x - y = \alpha + 1 \\ \alpha y = 2z = 2 \end{cases}$$

V.22. En el espacio afín E_3 y usando coordenadas cartesianas xyz, se consideran las siguientes rectas r y s (donde α y β son dados):

$$r\begin{cases} x + y - 3z + 2 = 0 \\ x - y - z + 4 = 0 \end{cases} \qquad s: \begin{cases} x = 2 + \alpha\lambda \\ y = 1 + \beta\lambda \\ z = 2\lambda \end{cases} \quad \lambda \in \mathbb{R}$$

Se pide: 1) valores de α y β para los que r y s se cortan; 2) valores de α y β para los que r y s son paralelas; 3) para $\alpha = 3$ y $\beta = 1$, hallar el plano paralelo a r y que pasa por s.

V.23. En el espacio afín E_3, se consideran las dos rectas r y s siguientes:

$$r: X = P + \lambda\vec{u} \quad (\lambda \in \mathbb{R})$$

y

$$s: X = Q + \mu\vec{v} \quad (\mu \in \mathbb{R})$$

Suponiendo que r y s se cruzan, hallar el lugar geométrico de los puntos medios de todos los segmentos que unen un punto de r con un punto de s.

V.24. Estudiar, en función del parámetro $\alpha \in \mathbb{R}$, la posi-

ción relativa de la recta r y el plano π siguientes o

$$r\begin{cases}x-y-z+1=0\\4x+\alpha y=1\end{cases} \qquad \pi:(\alpha+1)x+3y+2z=3$$

V.25. Estudiar, según los valores del parámetro $\alpha \in \mathbb{R}$, la posición relativa de los tres planos siguientes:

$$\begin{aligned}\alpha x+\ \ y+\ \ z&=1\\ x+\alpha^2 y+\ \ z&=\alpha\\ \alpha x+\ \ y+\alpha^3 z&=\alpha^2\end{aligned}$$

V.26. Estudiar, según los valores de los parámetros α y β la posición relativa de los tres planos siguientes:

$$\left.\begin{aligned}\pi_1:\alpha x+\ \ \ \ y+2z&=0\\ \pi_2:\alpha x+\ \ \beta y+2z&=1\\ \pi_3:\ \ x+(\beta-1)y+\ \ z&=0\end{aligned}\right\}$$

V.27. En el espacio euclídeo E_3 y usando coordenadas rectangulares xyz, se consideran las rectas r y s siguientes:

$$r\begin{cases}x+z=1\\ \alpha x+y+z=0\end{cases} \qquad s\begin{cases}2\alpha x+y+z=1\\ x+y+z+2=0\end{cases}$$

1. Hallar α de manera que r y s sean perpendiculares.
2. Para $\alpha=2$, hallar la proyección ortogonal de r sobre el segundo de los planos que definen a s.
3. Hallar el ángulo que forman la anterior proyección y s.

V.28. Sean $r=P+V(\vec{u})$ y $s=P+V(\vec{v})$ dos rectas, que pasan por un mismo punto P, del espacio euclídeo E_3. Se pide:

1. Las bisectrices del ángulo que forman r y s.
2. El lugar geométrico de los puntos X de E_3 que equidistan de r y s.

V.29. En el espacio euclídeo E_3 y usando coordenadas rectangulares $Oxyz$, se considera el plano $\pi: ax+by+cz+d=0$. Hallar la ecuación del lugar geométrico que engendran las rectas que pasan por el origen O y forman un ángulo θ con el plano π.

V.30. En el espacio euclídeo E_3 y utilizando coordenadas rectangulares $Ox_1x_2x_3$, se consideran tres puntos no alineados $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$ y $C(c_1, c_2, c_3)$, distintos de O; sea π un plano que se mueve paralelamente al ABC y sean $A'B'C'$ los puntos de intersección de π con las rectas OA, OB y OC, respectivamente. Por A', B' y C' se trazan planos perpendiculares a OA, OB y OC, respectivamente. Hallar el lugar geométrico descrito por el punto de intersección de estos tres planos.

V.31. En el espacio euclídeo E_3 y usando coordenadas rectangulares xyz, se consideran las rectas:

$$r\begin{cases}x=az+h\\ y=bz+k\end{cases} \qquad r'\begin{cases}x=a'z+h'\\ y=b'z+k'\end{cases}$$

Hallar la distancia entre ellas.

V.32. En el espacio euclídeo E_3 y usando coordenadas rectangulares $Oxyz$, se considera un triángulo ABC variable tal que sus vértices A, B y C se desplazan por los ejes Ox, Oy y Oz de manera que su área permanece constante e igual a S. Se pide:

1. Lugar geométrico de la proyección ortogonal del origen O sobre el plano ABC.
2. Lugar geométrico descrito por el baricentro del triángulo ABC.
3. Lugar geométrico descrito por el ortocentro del triángulo ABC.

V.33. En el espacio euclídeo se consideran los planos de ecuaciones euclídeas $\pi:\vec{a}\cdot\overrightarrow{PX}=0$ y $\tau:\vec{b}\cdot\overrightarrow{QX}=0$. Hallar la recta de π que pasa por P y es línea de máxima pendiente sobre el plano τ (se suponen π y τ no paralelos).

V.34. En el espacio euclídeo E_3, se consideran las rectas $r=P+V(\vec{u})$ y $s=Q+V(\vec{v})$, que se cruzan. Se considera una recta variable que corta a las r y s y es tal que forma con ellas ángulos iguales. Hallar el lugar geométrico descrito por el punto medio de la anterior recta variable.

V.35. Sea $OABC$ un tetraedro, del espacio euclídeo E_3. Si dos pares de aristas opuestas son perpendiculares, pruébese que el tercer par también está formado por aristas perpendiculares.

V.36. La «esfera» de un reloj de torre está situada en un plano π de manera que el extremo del minutero alcanza su posición más alta a las horas exactas. Sabiendo que dicha posición más alta es la del punto $P(5/3, 7/3, 5/3)$ y que el centro de la esfera

del reloj está situado en el punto $Q(1, 2, 1)$ (respecto de una referencia cartesiana rectangular $Oxyz$; el eje Oz vertical y hacia arriba). Se pide:

1. Ecuación del plano π.
2. Coordenadas del extremo del minutero a las 0^h15^m.
3. Para un observador situado en el punto $R(4, 1, 0)$, ¿cuál es el ángulo aparente que forman las posiciones del minutero a las 0^h y a las 0^h15^m?
4. Coordenadas del extremo del minutero a las $0^h\alpha^m$.
5. Ecuaciones (no paramétricas) de la circunferencia descrita por el extremo del minutero en el trascurso de una hora.

V.37. En el espacio afín ampliado $\bar{E}_3$ y utilizando coordenadas homogéneas x_0, x_1, x_2, x_3, se consideran tres planos

$$\Sigma a_i x_i = 0 \quad , \quad \Sigma b_i x_i = 0 \quad \text{y} \quad \Sigma c_i x_i = 0$$

($i = 0, 1, 2, 3$). Estudiar la posición relativa de dichos planos sin distinguir entre puntos propios e impropios (o sea, olvidándose del paralelismo).

V.38. En el espacio afín E_3 y respecto de una cierta referencia, se consideran los planos (en coordenadas homogeneas x_0, x_1, x_2, x_3):

$$\pi : a_0x_0 + a_1x_1 + a_2x_2 + a_3x_3 = 0$$

y

$$\tau : b_0x_0 + b_1x_1 + b_2x_2 + b_3x_3 = 0$$

En el estudio de la posición relativa de π y τ, recurriendo a las matrices

$$M = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \quad \text{y} \quad M' = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & b_3 \end{bmatrix}$$

alguien comete el error de intercambiar los subíndices 0 y 1 en las ecuaciones de π y τ. Con ello, ¿que significado tiene ahora cada una de las posibilidades siguientes?

1. rang $M = 2$.
2. rang $M = 1$ y rang $M' = 2$.
3. *rang* M = rang $M' = 1$.

V.39. En el espacio afín ampliado $\bar{E}_3$ y respecto de una referencia cartesiana, se consideran las rectas r y s que admiten por ecuaciones paramétricas (en coordenadas homogéneas x_0, x_1, x_2, x_3) a las:

$$r : (x_0, x_1, x_2, x_3) = \alpha(2, 1, 0, -1) + \beta(a, -1, 1, -2)$$
$$s : (x_0, x_1, x_2, x_3) = \alpha(3, 2, 1, -1) + \beta(2, -1, 2, -3)$$

(para $\alpha, \beta \in \mathbb{R}$).

1. Hallar el valor de α para el que r y s se cortan.
2. Para dicho valor de α, hallar la ecuación general del plano que contiene a r y s.
3. Hallar, en función de parámetros, las coordenadas de un punto genérico del infinito del plano de apartado anterior.

V.40. Sea E un espacio afín, asociado al espacio vectorial V, y considérese un subespacio vectorial U de V. Compruébese que la relación binaria $\sim$ definida, entre los puntos de E, mediante

$$[P \sim Q] \quad \Leftrightarrow \quad [\overrightarrow{PQ} \in U]$$

es una equivalencia. Hállense las clases de esta equivalencia.

V.41. Sea F un conjunto no vacío de puntos, de un espacio afín E (asociado al espacio vectorial V), que goza de la siguiente propiedad: si dos puntos distintos pertenecen a F, entonces también son de F todos los puntos de la recta que pasa por aquellos dos. Compruébese que F es una variedad lineal de F.

V.42. Sea F una variedad lineal, de un espacio afín. Hallar el lugar geométrico que describen los baricentros de cualquier familia finita de puntos de F.

V.43. Sea E un espacio afín asociado al espacio vectorial V. Si F_1 y F_2 son dos variedades lineales de F que tienen por direcciones a los subespacios U_1 y U_2 de V, compruébese que:

$$[F_1 \cap F_2 \neq \emptyset] \quad \Leftrightarrow$$
$$\Leftrightarrow \quad [(P_1 \in F_1, P_2 \in F_2) \quad \Rightarrow \quad \overrightarrow{P_1P_2} \in U_1 + U_2]$$

V.44. En un espacio afín E de dimensión n, dada una variedad afín $F = P + U$ de dimensión $h < n - 1$, hallar dos variedades lineales distintas F_1 y F_2 de dimensión h tales que $F_1 \cap F_2 = F$.

V.45. Sean $F_1 = P_1 + U_1$ y $F_2 = P_2 + U_2$ dos variedades lineales, de un cierto espacio afín, cuyas direcciones son U_1 y U_2. Si F_1 y F_2 son paralelas, hallar $F_1 \cap F_2$.

V.46. Sea $(A_0, A_1, ..., A_n)$ una referencia afín, de un espacio afín E de dimensión n. Sea $(B_0, B_2, ..., B_n)$ una nueva referencia afín de E; sean $(b_{0j}, b_{1j}, ..., b_{nj})$ las coordenadas baricéntricas de B_j, en la primera referencia. Si $(x'_0, x'_1, ..., x'_n)$ son las coordenadas baricéntricas de un punto $X \in E$ en la segunda referencia, hallar las coordenadas baricéntricas $(x_0, x_1, ..., x_n)$ de X en la primera referencia.

V.47. En un espacio afín E de dimensión n y respecto de una cierta referencia baricéntrica, se consideran los $h+1$, puntos independientes

$$P_0(p_{0i}),\ P_1(p_{1i}),\ ...,\ P_h(p_{hi}),\quad \text{donde } i = 0,\ 1,\ ...,\ n$$

Hallar, en función de las coordenadas de los puntos P_j y de cuantos parámetros sean necesarios, las coordenadas $(x_0, x_1, ..., x_n)$ de un punto genérico X de la variedad afín que engendran los $h+1$ puntos dados.

V.48. Si P_1, P_2 y P_3 son tres puntos alineados, de un espacio afín E, se llama razón simple de P_1, P_2, P_3 al número r, que se denota poniendo $r = (a_1 a_2 a_3)$, para el que $\overrightarrow{P_1P_3} = r\overrightarrow{P_1P_2}$. Se pide:

1. Si P_3 es el punto medio de los P_1 y P_2, hallar las razones simples

$$(P_1P_2P_3),\ (P_1P_3P_2),\ (P_2P_1P_3),$$
$$(P_3P_2P_1) \text{ y } (P_2P_3P_1)$$

2. Si $r = (P_1P_2P_3)$, hallar el baricentro de los puntos P_2 y P_3 afectados de los coeficientes r y -1.

V.49. Hallar la ecuación cartesiana del hiperplano, de un espacio afín E de dimensión 5, que pasa por los puntos P_0, P_1, P_2, P_3 y P_4 cuyas coordenadas, en una cierta referencia cartesiana son:

$$P_0(1,\ 3,\ -1,\ 4,\ 4),\quad P_1(0,\ 1,\ -1,\ 2,\ 2),$$
$$P_3(2,\ 0,\ 1,\ 2,\ 3),\quad P_4(1,\ -1,\ 1,\ 0,\ 1),$$
$$P_5(1,\ 3,\ -1,\ 1,\ 2)$$

V.50. En un espacio afín E de dimensión n, se consideran dos variedades lineales F_1 y F_2, de las que se sabe que, respecto de cierta referencia cartesiana:

1. La dirección de F_1 tiene por base al sistema de vectores $(\vec{v}_1, ..., \vec{v}_p)$, cuyas coordenadas son las columnas de una matriz dada M, de tamaño $n \times p$.
2. Las ecuaciones cartesianas de F_2, expresadas matricialmente, son $AX = B$ donde A es una matriz dada. de tamaños $h \times n$, X es la matriz columna $(n \times 1)$ de coordenadas.

Hallar la relación que debe verificarse entre A y M para que F_1 y F_2 sean paralelas.

V.51. En un espacio afín E, se consideran dos rectas r_1 y r_2 no paralelas. Hallar el lugar geométrico descrito por los puntos medios de todos los segmentos que unen un punto de r_1 con otro de r_2.

V.52. Sea E un espacio afín, asociado al espacio vectorial V. Considérense, en E, dos variedades afines $F_1 = P_1 + U_1$ y $F_2 = P_2 + U_2$, cuyas direcciones son U_1 y U_2; sea A el conjunto de los vectores $\overrightarrow{X_1X_2} \in V$, que se obtienen al recorrer X_1 y X_2 las variedades F_1 y F_2. Compruébese que A es una variedad lineal de V (considerado como espacio afín asociado a sí mismo).

V.53. Sean $F_1 = P_1 + U_1$ y $F_2 = P_2 + U_2$ dos variedades lineales de un espacio afín E, cuyas direcciones son U_1 y U_2; se supone que $F_1 \cap F_2 = \emptyset$. Se sabe que una tercera variedad lineal $F = P + U$, cuya dirección es U, es paralela a F_1 y a F_2. Analícese si F ha de ser paralela a las variedades lineales $F_1 \cap F_2$ y $F_1 + F_2$.

V.54. Sean dados h puntos P_1, P_2, ..., P_h, de un espacio afín euclídeo E, y sea P su baricentro, afectando a dichos puntos de los coeficientes c_1, c_2, ..., c_h (con $c_1 + c_2 + \cdots + c_h \neq 0$). Para cualquiera que sea $X \in E$, sea

$$f(X) = c_1\|\overrightarrow{XP_1}\|^2 + c_2\|\overrightarrow{XP_2}\|^2 + \cdots + c_h\|\overrightarrow{XP_h}\|^2$$

obtener $f(X)$ en función de $f(P)$ y de $\|\overrightarrow{XP}\|$.

V.55. Dadas dos rectas r_1 y r_2 que se cortan en un punto P, r_1 y r_2, de un espacio afín euclídeo E, se llaman bisectrices de r_1 y r_2 a las rectas que son coplanarias con ellas, pasan por P y forman ángulos iguales con r_1 y r_2. Hallar las bisectrices de r_1 y r_2.

V.56. Sean P, Q y R tres puntos, de un espacio afín euclídeo E. Pruébese que $d(P, Q) = d(P, R) + d(R, Q)$ sí y sólo si el punto R pertenece al segmento de extremos P y Q.

V.57. Sean P y Q dos puntos distintos, de un espacio afín

euclídeo E. Hallar el lugar geométrico que forman los puntos $X \in E$ que equidistan de P y Q.

V.58. En un espacio afín euclídeo de dimensión n, se considera la variedad lineal F que tiene por ecuación cartesiana (rectangular) a $AX = B$, donde A es matriz de tamaño $h \times n$ (la dimensión de F es $n - h$) y X es la matriz columna de coordenadas del punto genérico $X \in F$. ¿Qué son las filas de A? Describir cómo se determinaría la proyección ortogonal de un punto P sobre F.

V.59. En un espacio afín euclídeo E, se dan un punto P y el hiperplano H que pasa por el punto Q y es ortogonal al vector $\bar{a}$. Hallar la proyección ortogonal P_0 de P sobre H y el simétrico P' de P respecto de H.

V.60. En un espacio afín euclídeo E, se dan un punto P y la recta r que pasa por el punto Q y tiene la dirección del vector $\bar{u}$. Hallar la proyección ortogonal P_0 de P sobre r y el simétrico P' de P respecto de r.

V.61. En un espacio afín euclídeo E, se dan cuatro puntos fijos A, B, C y D. Hallar el lugar geométrico descrito por los puntos $X \in E$ tales que

$$\|\overrightarrow{XA} + 3\overrightarrow{XB}\| = \|\overrightarrow{XC} + 3\overrightarrow{XD}\|$$

V.62. En un espacio afín euclídeo E, se dan cuatro puntos fijos A, B, C y D. Hallar el lugar geométrico de los puntos $X \in E$ tales que

$$\overrightarrow{XA} \cdot (\overrightarrow{XB} + \overrightarrow{XC} + \overrightarrow{XD}) = \alpha \quad (\alpha \in \mathbb{R} \text{ dado})$$

V.63. En un espacio afín euclídeo E, se consideran $h + 1$ puntos fijos $A_0, A_1, \ldots, A_h$. Hallar el lugar geométrico descrito por los puntos $X \in E$ tales que

$$\left\| \sum_{i=1}^{h} c_i \overrightarrow{XA_i} \right\| = \left| \sum_{i=1}^{h} c_i \right| \|\overrightarrow{XA_0}\|$$

donde $c_1, c_2, \ldots, c_n$ son números reales que tienen suma no nula.

V.64. En un espacio afín euclídeo E, se considera una esfera S de radio ρ y un punto P situado a distancia d del centro C de S. Por P se traza una recta variable que corta a S en X_1 y X_2. Si $[PX]$ y $[PX_2]$ denotan las distancias orientadas de P a X_1 y X_2 (distancias con igual o distinto signo según que X_1 y X_2 estén o igual o distinto lado de P), pruébese que $[PX_1][PX_2] = d^2 - \rho^2$ (constante que se llama *potencia* de P respecto de S).

V.65. En un espacio afín euclídeo E, se consideran dos esferas S_1 y S_2. Hallar el lugar geométrico de los puntos que tienen la misma potencia (véase el ejercicio anterior) respecto de ambas esferas. Este lugar es un hiperplano, que se llama *hiperplano radical* de las esferas dadas.

SOLUCIONES

V.1. La recta P_1P_2 $(X = P_1 + \lambda\overrightarrow{P_1P_2})$ ha de cortar a r para $O < \lambda < 1$;

$$\text{signo } (a\alpha_1 + b\beta_1 + c) \neq \text{signo } (a\alpha_2 + b\beta_1 + c)$$

V.2. Tómese la referencia cartesiana que tiene su origen en O, que tiene por ejes a las rectas OAA' y OBB' y cuyos vectores son tales que C tiene coordenadas $(1, 1)$; $A(a, 0)$, $A'(a', 0)$, $B(0, b)$, $B'(0, b')$, $C'(h, h)$.

V.3. Tómese la referencia $(A_1; \overrightarrow{A_1A_2}, \overrightarrow{A_1A_3})$. Si las coordenadas i-ésimas de tres puntos P, Q y R son p_i, q_i y r_i, entonces $(PQR) = (r_i - p_i)/(q_i - p_i)$. Sean $B_3(a, 0)$ y $B_2(0, b)$ con lo que

$$B_1[(ab - a)/(b - a),\quad (ba - b)/(a - b)]$$

$$(B_1A_2A_3) = \frac{a(1-b)}{b(1-a)}$$

$$(B_2A_3A_1) = \frac{b}{b-1}$$

$$(B_3A_1A_2) = \frac{a-1}{a}$$

V.4. Tómese la referencia $(A_1; \overrightarrow{A_1A_2}, \overrightarrow{A_1A_3})$; en ella $B_1(\alpha, 1 - \alpha)$, $B_2(0, b)$ y $B_3(a, 0)$

$$(B_1A_2A_3)(B_2A_3A_1)(B_3A_1A_2) = \frac{\alpha b(a-1)}{(1-\alpha)(1-b)a} = H$$

Las rectas $x + y/b = 1$, $x/a + y = 1$ y $(\alpha - 1)x = \alpha y$ se cortan en un punto si $H = -1$.

V.5. Sean r_1, r_2, r_3 y r_4 los lados; tómese referencia cartesiana de ejes r_1 y r_2 y en la que las coordenadas de $r_1 \cap r_3$ y de $r_2 \cap r_4$ son $(1, 0)$ y $(0, 1)$; sean $(a, 0)$ y $(0, b)$ las coordenadas de $r_1 \cap r_4$ y de $r_2 \cap r_3$. Los puntos medios son $(1/2, 1/2)$, $(a/2, b/2)$ y $[(ab - a)/(2ab - 2), (ab - b)/(2ab - 2)]$, que están alineados.

V.6. El punto medio de AC es el $B + (3/2)\overrightarrow{BG}$, o sea $(4, 5)$; luego $C(15, 7)$.

V.7. Tómense a r y s como ejes y a B como punto unidad, $B(1, 1)$ y $A(2,2)$. Sea $P(\alpha, 0)$ con $\alpha \in \mathbb{R}$; $Q[0, \alpha/(\alpha - 1)]$,

$$AP: 2x + (\alpha - 2)y = 2\alpha \quad QM: y - x = \alpha/(\alpha - 1)$$

Eliminando α entre estas dos ecuaciones, resulta $(y - 2x)(y - x) = 0$, que son dos rectas; la $y - x = 0$ es la OA.

V.8. Tómese referencia rectangular en la que $A(\alpha, 0)$, $B(\alpha', 0)$ y $C(0, \beta)$; las alturas se cortan en el punto $(0, -\alpha\alpha'/\beta)$.

V.9. Ortocentro (intersección de las alturas) $(1, 2)$; el baricentro (intersección de las medianas) $(4/3, 2)$; el circuncentro (intersección de las mediatrices) $(3/2, 2)$.

V.10. $3x + 4y + 7 = 0$. $3x + 4y - 3 = 0$, $4x - 3y + 16 = 0$. $4x - 3y + 6 = 0$.

V.11. $A(a, 2a)$, $B(b, -2b)$;

$x = \frac{1}{3}(a + b)$, $y = \frac{2}{3}(a - b)$;

$9 = |4ab|$.

Eliminando a y b entre las tres últimas ecuaciones: $4x^2 - y^2 = \pm 4$ (dos hipérbolas).

V.12. $$x = a + \frac{\operatorname{sen}(\theta - \alpha)}{\operatorname{sen}\theta}x' + \frac{\operatorname{sen}(\theta - \beta)}{\operatorname{sen}\theta}y'$$

$$y = b + \frac{\operatorname{sen}\alpha}{\operatorname{sen}\theta}x' + \frac{\operatorname{sen}\beta}{\operatorname{sen}\theta}y'$$

V.13. 1. $b^2 > a$.
2. $y^2 = x$ (parábola).
3. La pendiente es $m = (-1/2)\alpha$ y la perpendicularidad equivale a $m_1m_2 = -1$ o sea $\alpha_1\alpha_2 = -1/4$; de la ecuación de α sale $\alpha_1\alpha_2 = x$, luego el lugar: $x = -1/4$.

V.14. Tómense por ejes (rectangulares) Ox y Oy al lado BC y a su altura; $A(0, a)$, $B(b, 0)$ y $C(-b, 0)$; $Q(\alpha, 0)$, $\alpha \in \mathbb{R}$; $PQ: y = m(x - \alpha)$, $m \in \mathbb{R}$; $b^2m + a = 0$; la perpendicular $b^2x - \alpha(ay + b^2) = 0$, que es un haz, luego pasan por un punto.

V.15. Sea P el punto de intersección de $ax + by + c = 0$ y $\alpha x + \beta y = 0$; sea Q el de intersección de $\alpha x + \beta y + \gamma = 0$ con $ax + by = 0$. El área es $4\|\overrightarrow{OP} \wedge \overrightarrow{OQ}\|$, donde O es el origen, que vale $4|c\gamma|:|a\beta - b\alpha|$.

V.16. Tomando por ejes rectangulares (Oxy) a las bisectrices de r y s, éstas tendrán ecuaciones $y = \pm mx$ ($m \in \mathbb{R}$ fijo). La suma de las distancias de $P(x, y)$ a r y s vale

$$\frac{|y - mx| + |y + mx|}{\sqrt{1 + m^2}}$$

igualándolo a k, se obtienen los cuatro lados del rectángulo de vértices

$$\left(\pm\frac{k\sqrt{1+m^2}}{2m}, 0\right) \quad \text{y} \quad \left(0, \pm\frac{k\sqrt{1+m^2}}{2}\right)$$

V.17. 1. $y-a=m(x-a)$ será tangente si dista r de O;

$$m^2[(\alpha-a)^2-r^2]-2(\alpha-a)(\beta-b)m+$$
$$+(\beta-b)^2-r^2=0$$

2. Sean P_1 y P_2 las intersecciones de OO' con las tangentes exteriores e interiores;

$$\|\overrightarrow{P_1O_1}\|=(rd)/(r'-r) \text{ y } \|\overrightarrow{P_2O_1}\|=(rd)/(r'+r)$$

Sean $\pm\theta_1$ y $\pm\theta_2$ los ángulos de OO' con las tangentes exteriores e interiores;

$$\text{sen } \theta_1=(r'-r)/d \quad , \quad \text{sen } \theta_2=(r'+r)/d$$

Las tangentes son:

$$\text{ext.:}\begin{cases} x=\alpha-\dfrac{rd}{r'-r}\cos\varphi-\lambda\cos(\varphi\pm\theta_1) \\ y=\beta-\dfrac{rd}{r'-r}\text{sen }\varphi+\lambda\text{ sen }(\varphi\pm\theta_1)\end{cases}$$

$$\text{int.:}\begin{cases} x=\alpha-\dfrac{rd}{r+r'}\cos\varphi-\lambda\cos(\varphi\pm\theta_2) \\ y=\beta-\dfrac{rd}{r+r'}\text{sen }\varphi+\lambda\text{ sen }(\varphi\pm\theta_2)\end{cases}$$

V.18. 1. Si $X=C+QX'$, es $X'=Q^{-1}X-Q^{-1}C$; $x'=2x+2y-2z$, $y'=-9x-6y+7z-5$, $z'=-x-y+z$.

2. Resolviendo $X=C+QX$, se obtiene un solo punto $X(1,-1,1)$.

V.19. 1. Es la intersección de los planos que pasan por P y r y por P y s;

$$(x,y,z)=(-1,1,2)+\lambda(1,2,3)$$

2. Intersección del plano que pasa por P y r y del plano paralelo a π por P;

$$(x,y,z)=(-1,1,2)+\lambda(0,1,1)$$

V.20. $(-2,-1,0)$, $(0,-1,-1)$, $(0,-3/2,1/2)$ y $(-2/3,2/3,-4/3)$.

V.21. Si $\alpha=-2$ las rectas son paralelas y distintas; si $\alpha=-7$, las rectas se cortan en un punto; si $\alpha\neq-2,-7$, las rectas se cruzan.

V.22. 1. α cualquiera, $\beta=2$.
2. $\alpha=4$ y $\beta=2$.
3. $x-y-z=1$.

V.23. El punto medio de $P+\lambda\bar{u}$ y $Q+\mu\bar{v}$ es

$$O-(\lambda/2)\bar{u}+(\mu/2)\bar{v}$$

donde $O=P+(1/2)\overrightarrow{PQ}$ es el punto medio de las P y Q. El lugar es el plano paralelo a r y a s por O.

V.24. Si $\alpha=-4$, es paralela a π y no está incluida en él; si $\alpha=1$, r está incluida en π; si $\alpha\neq-4,1$, entonces r corta a π (en un punto).

V.25. Si $\alpha=1$, los tres planos coinciden; si $\alpha\neq1$, los tres planos se cortan en un punto (forman triedro).

V.26. Si $\alpha\neq2$ y $\beta\neq1$, la intersección es un punto (forman triedro); si $\alpha\neq2$ y $\beta=1$, la intersección es vacía y π_1 y π_2 son paralelos; si $\alpha=2$ y $\beta=3/2$ la intersección es una recta y π_1 y π_3 coinciden; si $\alpha=2$ y $\beta=1$, la intersección es vacía y π_1 y π_2 son paralelos; si $\alpha=2$ y $\beta=2$, la intersección es vacía y π_2 y π_3 son paralelos; si $\alpha=2$ y $\beta\neq\neq3/2$. 1, 2, la intersección es vacía y los planos forman un prisma.

V.27. 1. $(1,1-\alpha,-1)$ y $(0,2\alpha-1,1-2\alpha)$ han de ser ortogonales; $\alpha=2$ y $\alpha=1/2$.

2. La intersección de $x+y+z+2=0$ con el plano del haz de r perpendicular al anterior, que dá:

$$\begin{cases} x+y+z+2=0 \\ y-z+2=0\end{cases}$$

3. Son perpendiculares, pues lo eran r y s.

V.28. Sean $\bar{u}_1$ y $\bar{v}_1$ los vectores unitarios de las direcciónes de $\bar{u}$ y $\bar{v}$.

1. $P+\mathcal{V}(\bar{u}_1+\bar{v}_1)$ y $P+\mathcal{V}(\bar{u}_1-\bar{v}_1)$.
2. $d(X,r)^2=\|\overrightarrow{PX}\|^2-(\overrightarrow{PX}\cdot\bar{u}_1)^2$,
$d(X,s)=\|\overrightarrow{PX}\|^2-(\overrightarrow{PX}\cdot\bar{v}_1)^2$;
$\overrightarrow{PX}\cdot\bar{u}_1=\pm\overrightarrow{PX}\cdot\bar{v}_1$;
el lugar lo forman los dos planos

$$(\bar{u}_1+\bar{v}_1)\cdot\overrightarrow{PX}=0 \quad \text{y} \quad (\bar{u}_1-\bar{v}_1)\cdot\overrightarrow{PX}=0$$

V.29. $(ax + by + cz)^2 = (a^2 + b^2 + c^2)(x^2 + y^2 + z^2)\cos^2\theta$.

V.30. Los planos son (para $\lambda \in \mathbb{R}$; las sumas para $i = 1, 2, 3$):

$$\Sigma a_i x_i = \lambda \Sigma a^2, \quad \Sigma b_i x_i = \lambda \Sigma b_i^2 \quad \text{y} \quad \Sigma c_i x_i = \lambda \Sigma c_i^2$$

eliminando λ, se obtiene el lugar:

$$\frac{\Sigma a_i x_i}{\Sigma a_i^2} = \frac{\Sigma b_i x_i}{\Sigma b_i^2} = \frac{\Sigma c_i x_i}{\Sigma c_i^2} \quad \text{(una recta)}$$

V.31. Un punto y un vector de r son $P(h, k, 0)$ y $\bar{u}(a, b, 1)$; un punto y un vector de r' son $P'(h', k', 0)$ y $\bar{u}'(a', b', 1)$. La distancia vale:

$$d = \frac{|\overrightarrow{PP'} \cdot (\bar{u} \wedge \bar{u}')|}{\cdot \|\bar{u} \wedge \bar{u}'\|} =$$

$$= \frac{|(a - a')(k - k') - (b - b')(h - h')|}{\sqrt{(a - a')^2 + (b - b')^2 + (ab' - a'b)^2}}$$

V.32. $A(a, 0, 0)$, $B(0, b, 0)$ y $C(0, 0, c)$; el área del triángulo vale $\|\overrightarrow{AB} \wedge \overrightarrow{AC}\|/4$, luego

$$(ab)^2 + (bc)^2 + (ca)^2 = 4S \qquad [1]$$

1. Se eliminan a, b y c entre [1],

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1 \quad \text{y} \quad ax = by = cz$$

y se tiene:

$$(x^2 + y^2 + z^2)^4 \left(\frac{1}{(xy)^2} + \frac{1}{(yz)^2} + \frac{1}{(zx)^2} \right) = 4S^2$$

2. Se elimina a, b y c entre [1] y $x = a/3$, $y = b/3$, $z = c/3$ y se tiene:

$$81(x^2y^2 + y^2z^2 + z^2x^2) = 4S^2$$

3. El ortocentro es la proyección del origen sobre ABC, luego el lugar es el mismo que en 1.

V.33. La línea de máxima pendiente de π sobre τ es la recta de π que es ortogonal a la intersección $\pi \cap \tau$, luego ella es:

$$P + \mathbb{V}[\bar{a} \wedge (\bar{a} \wedge \bar{b})]$$

V.34. Sean $\bar{u}_0$ y $\bar{v}_0$ vectores unitarios de las direcciones de $\bar{u}$ y $\bar{v}$; sea θ el ángulo de r y s. Si $X \in r$ e $Y \in s$,

$$\overrightarrow{XY} = \overrightarrow{PQ} - \lambda\bar{u} + \mu\bar{v}$$

debe ser tal que

$$\overrightarrow{XY} \cdot \bar{u}_0 = \pm\overrightarrow{XY} \cdot \bar{v}_0$$

luego $\lambda \pm \mu = 4H$ siendo

$$4H = [\overrightarrow{PQ} \cdot (\bar{u}_0 \mp \bar{v}_0)]/(1 \mp \cos\theta)$$

se puede poner $\lambda = 2H + 2t$ y $\mu = \pm 2H \mp 2t$ con $t \in \mathbb{R}$. El punto medio buscado es

$$M = P + \lambda\overrightarrow{PQ} + \tfrac{1}{2}\overrightarrow{XY} =$$
$$= O + H(\bar{u}_0 \pm \bar{v}_0) + t(\bar{u}_0 \mp \bar{v}_0)$$

donde O es el punto medio de los P y Q. El lugar es las dos rectas que pasan por

$$M_0 = O + H(\bar{u}_0 \pm \bar{v}_0)$$

y tiene la dirección del vector $\bar{u}_0 \mp \bar{v}_0$.

V.35. Se supone $\overrightarrow{OA} \cdot \overrightarrow{BC} = 0$ y $\overrightarrow{OB} \cdot \overrightarrow{CA} = 0$; hay que probar que $\overrightarrow{OC} \cdot \overrightarrow{AB} = 0$;

$$0 = \overrightarrow{OA} \cdot \overrightarrow{BC} = \overrightarrow{OA} \cdot (\overrightarrow{OC} - \overrightarrow{OB})$$

y

$$0 = \overrightarrow{OB} \cdot \overrightarrow{CA} = \overrightarrow{OB} \cdot (\overrightarrow{OA} - \overrightarrow{OC})$$

restando éstas, se obtiene

$$0 = \overrightarrow{OC} \cdot (\overrightarrow{OB} - \overrightarrow{OA}) = \overrightarrow{OC} \cdot \overrightarrow{AB}$$

V.36. Longitud del minutero $= \|\bar{v}\| = \|\overrightarrow{PQ}\| = 1$.

1. La dirección de $\bar{v}$ es la de máxima pendiente de π, luego es perpendicular a $\bar{w} = (4, 2, -5)$; es $\pi: 4x - 2y - 5 = 3$.
2. El minutero ocupa la posición del vector $\bar{u}$ unitario de $\bar{v} \wedge \bar{w}$; el extremo es

$$(1 - 1/\sqrt{5},\ 2 + 2/\sqrt{5},\ 1)$$

3. Es el ángulo de $\mathbb{V}(\overrightarrow{RQ}, \bar{v})$ con $\mathbb{V}(\overrightarrow{RQ}, \bar{u})$ y vale

$$\text{arc cos}\,(1/2\sqrt{3}) \simeq 73{,}22°.$$

4. $\theta = (\pi\alpha)/30$; $Q + (\cos\theta)\bar{v} + (\text{sen}\,\theta)\bar{u}$;

$$x = 1 + \frac{2}{3}\cos\theta - \frac{1}{\sqrt{5}}\,\text{sen}\,\theta$$

$$y = 2 + \frac{1}{3}\cos\theta + \frac{2}{\sqrt{5}}\,\text{sen}\,\theta$$

$$z = 1 + \frac{2}{3}\cos\theta$$

5. $(x-1)^2 + (y-2)^2 + (z-1)^2 = 1$,
$4x + 2y - 5z = 3$

V.37. Sea M la matriz (de tamaño 3×4) cuyas filas son los coeficientes de cada uno de los tres planos. Si rang $M = 3$, los planos se cortan en un punto (propio o impropio); si rang $M = 2$, los planos se cortan a lo largo de una recta (propia o impropia); si rang $M = 1$, los planos coinciden.

V.38. 1. π y τ se cortan en una recta distinta de la $x_1 = 0$ (puede ser, en particular, la recta del infinito).
2. π y τ se cortan en la recta $x_1 = 0$.
3. π y τ coinciden.

V.39. 1. Ha de ser nulo el determinante cuyas filas son $(2, 1, 0, -1)$, $(a, -1, 1, -2)$, $(3, 2, 1, -1)$ y $(2, -1, 2, -3)$; $a = 1$.
2. Igualando a cero el determinante cuyas filas son (x_0, x_1, x_2, x_3) y tres de las cuatro anteriores, se tiene $x_0 - x_1 + x_3 = 0$.
3. $x_0 = 0$, $x_1 = \lambda$, $x_2 = \mu$, $x_3 = \lambda$.

V.40. Es reflexiva, simétrica y transitiva. La clase de representante $P \in E$ es la variedad lineal $P + U$.

V.41. Sea $P \in F$ y $U = \{\bar{u}_x = \overrightarrow{PX}/X \in F\}$; el conjunto U es subespacio vectorial de V: dado $\bar{v} = \lambda\bar{u}_x + \mu\bar{u}_y$, como $X' = P - \lambda\bar{u}_x$ e $Y' = P + (\mu/2)\bar{u}_y$ son de F y, por serlo éstos, también lo es

$$Z = X' + 2[(\mu/2)\bar{u}_y + \lambda\bar{u}_x] = P + \lambda\bar{u}_x + \mu\bar{u}_y$$

luego $\lambda\bar{u}_x + \mu\bar{u}_y \in U$.

V.42. La propia variedad F.

V.43. $\Rightarrow$) $\exists P \in F_1 \cap F_2$; $\overrightarrow{PP_1} \in U_1$, $\overrightarrow{PP_2} \in U_2$,
$\overrightarrow{P_1P_2} = \overrightarrow{PP_2} - \overrightarrow{PP_1} \in U_1 + U_2$

$\Leftarrow$) $\overrightarrow{P_1P_2} = \bar{u}_1 + \bar{u}_2$, $\bar{u}_1 \in U_1$ y $\bar{u}_2 \in U_2$;
$P = P_1 + \bar{u}_1 = P_2 - \bar{u}_2 \in F_1 \cap F_2$

V.44. Sean $\bar{v}_1$ y $\bar{v}_2$ tales que

$$\dim[U + V(\bar{v}_1, \bar{v}_2)] = h + 2 \leqslant n$$

si $U_1 = U + V(\bar{v}_1)$ y $U_2 = U + V(\bar{v}_2)$, tómese

$$F_1 = P + U_1 \quad \text{y} \quad F_2 = P + U_2$$

V.45. $F_1 \cap F_2 = F_1$ o F_2 o $\varnothing$.

V.46. $x_i = b_{i0}x'_0 + b_{i1}x'_1 + \cdots + b_{in}x'_n$ $(i = 0, 2, \ldots, n)$.

V.47. $\overrightarrow{P_0X} = \lambda_1\overrightarrow{P_0P_1} + \cdots + \lambda_h\overrightarrow{P_0P_h}$, sea

$$\lambda_0 = 1 - \lambda_1 - \cdots - \lambda_h$$

$$\overrightarrow{OX} = \lambda_0\overrightarrow{OP_0} + \lambda_1\overrightarrow{OP_1} + \cdots + \lambda_h\overrightarrow{OP_h}$$

$$x_i = \lambda_0 p_{0i} + \lambda_1 p_{1i} + \cdots + \lambda_h p_{hi}$$

con $\lambda_0 + \lambda_1 + \cdots + \lambda_h = 0$ $(i = 0, 1, \ldots, n)$.

V.48. 1. $1/2$, 2, $1/2$, -1, 2.
2. P_1.

V.49. $4x_1 - x_3 + 4x_4 - 6x_5 + 3 = 0$.

V.50. $A \cdot M = O$.

V.51. $r_1 : X_1 = P_1 + \lambda\bar{u}_1$, $r_2 : X_2 = P_2 + \lambda\bar{u}_2$;
$X = (P + 1/2\overrightarrow{PQ}) - (\lambda/2)\bar{u} + \mu\bar{v}$

Plano paralelo a r_1 y a r_2 que pasa por el punto medio de P y Q.

V.52. $A = \overrightarrow{P_1P_2} + (U_1 + U_2)$ («pasa» por $\overrightarrow{P_1P_2}$ y tiene por dirección a $U_1 + U_2$).

V.53. Las direcciones de $F_1 \cap F_2$ y $F_1 + F_2$ son $U_1 \cap U_2$ y $U_1 + U_2$. Puede ocurrir que $U_1, U_2 \subset U$, que $U \subset U_1$, U_2 o que $U_1 \subset U \subset U_2$; en los tres casos, $U_1 \cap U_2$ y $U_1 + U_2$ incluyen o están incluidos en U, luego F es paralela a $F_1 \cap F_2$ y a $F_1 + F_2$.

V.54. Poniendo $\overrightarrow{XP_i} = \overrightarrow{XP} + \overrightarrow{PP_i}$, se obtiene

$$f(X) = f(P) + (c_1 + c_2 + \cdots + c_h)\|\overrightarrow{XP}\|^2$$

V.55. Si $V(\bar{u}_1)$ y $V(\bar{u}_2)$ son las direcciones de r_1 y r_2, las bisectrices son

$$P + V\left(\frac{\bar{u}_1}{\|\bar{u}_1\|} \pm \frac{\bar{u}_2}{\|\bar{u}_2\|}\right)$$

V.56. Si $R \in [P, Q]$, póngase $\overrightarrow{PR} = r\overrightarrow{PQ}$ con $0 < r < 1$. Si se verifica la igualdad, es $\|\bar{u} + \bar{v}\| = \|\bar{u}\| + \|\bar{v}\|$, con $\bar{u} = \overrightarrow{PR}$ y $\bar{v} = \overrightarrow{RQ}$, o sea $\bar{u} \cdot \bar{v} = 0$.

V.57. Hiperplano ortogonal a $\overrightarrow{PQ}$ por su punto medio;

$$\{X \in E / 2\overrightarrow{PX} \cdot \overrightarrow{PQ} = \|\overrightarrow{PQ}\|^2\}$$

V.58. Las filas de A son las filas de coordenadas de un sistema generador del subespacio ortogonal a F. Ha de ser $X = P + \bar{v}$, con $\bar{v}$ ortogonal a F, $\bar{v} = A'\Lambda$ (Λ columna de h parámetros). Como $X \in F$, ha de ser $A(P + A'\Lambda) = B$, de aquí se obtendría Λ.

V.59. $\overrightarrow{PP_0} = \lambda\bar{a}$ y $\overrightarrow{QP_0} \cdot \bar{a} = 0$;

$$P_0 = P - \frac{\overrightarrow{QP} \cdot a}{\|\bar{a}\|^2}\bar{a} \quad , \quad P' = P - 2\frac{\overrightarrow{QP} \cdot a}{\|\bar{a}\|^2}\bar{a}$$

V.60. $\overrightarrow{PP_0} \cdot \bar{u} = 0$ y $\overrightarrow{QP_0} = \lambda\bar{u}$;

$$P_0 = Q - \frac{\overrightarrow{PQ} \cdot \bar{u}}{\|\bar{u}\|^2}\bar{u} \ , \ P' = P + 2\overrightarrow{PQ} - 2\frac{\overrightarrow{PQ} \cdot \bar{u}}{\|\bar{u}\|^2}\bar{u}$$

V.61. $\overrightarrow{XA} + 3\overrightarrow{XB} = 4\overrightarrow{XA'}$ y $\overrightarrow{XC} + 3\overrightarrow{XD} = 4\overrightarrow{XC'}$, donde A' y C' son los dos puntos $A' = B + (1/4)\overrightarrow{BA}$ y $C' = D + (1/4)\overrightarrow{DC}$. El lugar es el hiperplano ortogonal a $\overrightarrow{A'C'}$ por su punto medio.

V.62. Sea G el baricentro (con coeficientes unidad) de B, C y D; sea O el baricentro de G y A;

$$\overrightarrow{XA} \cdot (3\overrightarrow{XGR}) = \alpha \quad ; \quad \|\overrightarrow{XO}\|^2 = \|\overrightarrow{OA}\|^2 + \alpha/3$$

Esfera de centro en O y radio $\rho = [\|\overrightarrow{OA}\|^2 + \alpha/3]^{1/2}$.

V.63. Sea $\overset{\frown}{G}$ el baricentro (con coeficientes unidad) de los puntos $A_1, \ldots, A_h$;

$$\left\| \left| \sum_{i=1}^{h} c_i(\overrightarrow{XG} + \overrightarrow{GA_i}) \right| \right| = \left| \sum_{i=1}^{h} c_i \right| \|\overrightarrow{XA_0}\|$$

$\|\overrightarrow{XG}\| = \|\overrightarrow{XA_0}\|$ hiperplano ortogonal a $\overrightarrow{GA_0}$ por su punto medio.

V.64. $\overrightarrow{PX_i} = \lambda_i\bar{u}$ ($\bar{u}$ unitario); $\|\overrightarrow{CP} + \overrightarrow{PX_i}\| = \rho$;
$\lambda^2 + 2\lambda\overrightarrow{CP} \cdot \bar{u} + (d^2 - \rho^2) = 0$;
$[PX_1][PX_2] = \lambda_1\lambda_2 = d^2 - \rho^2$.

V.65. Sean C_1, C_2, ρ_1 y ρ_2 los centros y los radios;

$$\|\overrightarrow{C_1X}\|^2 - \rho_1^2 = \|\overrightarrow{C_2X}\| - \rho_2^2$$

Poniendo

$$\overrightarrow{C_1X} = \lambda\overrightarrow{C_1C_2} + \bar{u} \quad \text{y} \quad \overrightarrow{C_2X} = (\lambda - 1)\overrightarrow{C_1C_2} + \bar{u}$$

con $\bar{u} \perp \overrightarrow{C_1C_2}$, es

$$\|\overrightarrow{C_1X}\|^2 = \lambda^2 d^2 + \|\bar{u}\|^2$$

y

$$\|\overrightarrow{C_2X}\|^2 = (\lambda - 1)^2 d^2 + \|\bar{u}\|^2$$

El lugar es el hiperplano ortogonal a $\overrightarrow{C_1C_2}$ por el punto $C = C_1 + \lambda\overrightarrow{C_1C_2}$, donde

$$\lambda = (\rho_1^2 - \rho_2^2 + d^2)/(2d^2)$$

PARTE VI

AQUÍ, el autor cree que ha de confesar que ha tenido dudas, y no pocas, acerca de cómo encauzar esta última parte de la obra. El autor, dando tumbos y vaivenes, llegó incluso a pensar que debía omitirla, cosa que, a poco, desechó de sí, como el que huye del diablo.

El autor ha estimado, a la postre, que el texto quedaría asaz incompleto si, en él, no se hablase de los lugares geométricos de segundo grado u orden, que de ambos modos se dice. A juicio del autor, dichas curvas y superficies cuadráticas han de interesar a muchos de los que, por gusto o por necesidad, están ocupados en esto de iniciarse en el conocimiento del álgebra y de la geometría.

El autor se barrunta que el interés, de los hipotéticos lectores, por las cónicas y las cuádricas, que así se llama a las ya referidas curvas y superficies, no va demasiado lejos y, consecuente con ello, no se ha adentrado en exceso en la presentación que hace de aquéllas.

El autor piensa que su exposición, sobre el asunto de las cónicas y las cuádricas, no es desatinada, que guarda un razonable equilibrio entre situaciones extremas. El autor desea que así sea, confía en no haber errado el tiro, pero cabe la posibilidad de haber hecho un traje que a unos les quede corto y que para otros sea largo. El autor pide al cielo no haberse equivocado.

13. Cónicas y cuádricas: estudio particular.
14. Cónicas y cuádricas: análisis afín y métrico. Ejercicios y problemas.

CAPITULO

13

Cónicas y cuádricas: Estudio particular

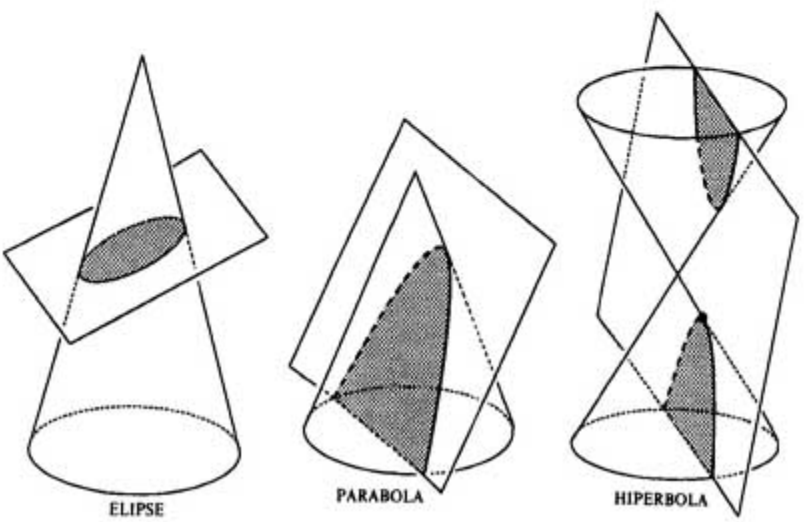
Dedicamos estos dos últimos capítulos al estudio de las cónicas y de las cuádricas, esto es de las curvas y de las superficies que admiten ecuación de segundo grado, por lo que se llaman también curvas y superficies de segundo orden. En lugar de acometer, como es usual, primero el análisis de las cónicas y, acabado éste, empezar con el de las cuádricas, procederemos, por el contrario en paralelo, avanzando alternativamente en el examen de las unas y de las otras. Así, agruparemos las cuestiones que, por ser similares para ambas, admiten análogo tratamiento.

Nos ocuparemos primero de las cuestiones particulares y sencillas; la complejidad irá en aumento, para acabar con aquellos asuntos más generales y difíciles. Empezaremos con el «análisis particular de las cónicas y las cuádricas» que, parcialmente al menos, nos debe ser ya conocido de estudios anteriores. Más tarde, en el próximo capítulo, abordaremos el estudio métrico de las cónicas y de las cuádricas, esto es, las analizaremos con generalidad desde el punto de vista euclídeo.

α ESTUDIO PARTICULAR DE LAS CONICAS

Empezaremos presentando a las elipses, a las hipérbolas y a las parábolas, por separado, definiéndolas como lugares geométricos del plano euclídeo E_2. Obtendremos sus «ecuaciones reducidas», es decir, ecuaciones respecto de unas referencias *ad hoc* que las hacen especialmente sencillas.

Estudiaremos sus propiedades más sobresalientes, señalando sus parecidos y sus diferencias. Terminaremos dando un tratamiento unificado a los tres tipos de cónicas, o secciones cónicas, valiéndonos de sus «ecuaciones focales».



13.1. LAS TRES CONICAS

LA ELIPSE; ECUACION REDUCIDA

[278]

En el plano euclídeo E_2, se llama *elipse* que tiene por *focos* a los puntos F y F' (situados a distancia $\overline{FF'} = 2c$) y cuya *constante* es $2a \in \mathbb{R}$ (siendo $a > c$), al lugar geométrico de los puntos $X \in E_2$ tales que $\overline{XF} + \overline{XF'} = 2a$. Se llaman *ejes* de la elipse, por ser sus ejes de simetría (ortogonal), a la recta FF' (eje focal o mayor) y a su mediatriz (eje secundario o menor); su punto de intersección, O, es su *centro* de simetría. Los puntos de la elipse que están en sus ejes se llaman *vértices*; hay dos, A y A', en el eje mayor y otros dos B y B' en el eje menor. Se verifica que

$$\overline{OA} = \overline{OA'} = a,\ \overline{FB} = \overline{F'B} = \overline{FB'} = \overline{F'B'} = a,\ \overline{OB} = \overline{OB'} = b$$

siendo $b > 0$ tal que se verifica la relación:

$$a^2 = b^2 + c^2 \quad \left(\begin{array}{l} a \text{ y } b \text{ se llaman longitudinales de los} \\ \text{semiejes; } c \text{ es la semidistancia focal} \end{array}\right)$$

La elipse referida a sus ejes mayor (Ox) y menor (Oy) admite por ecuación a:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{(ecuación reducida) o}$$

$$\begin{cases} x = a \cos \theta \\ y = b \operatorname{sen} \theta \end{cases} \quad \text{para} \quad 0 \leq \theta < 2\pi \quad \text{(ecuaciones paramétricas)}$$

La elipse está situada entre las circunferencias de su mismo centro O que tienen radios a y b.

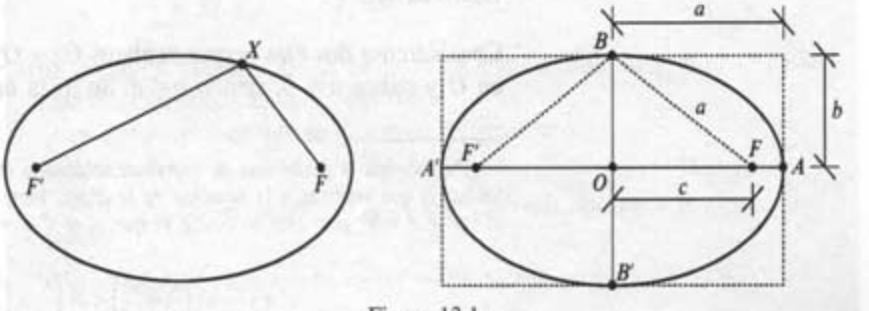


Figura 13.1.

DEMOSTRACION

De su definición se desprende, obviamente, que la elipse es simétrica respecto de la recta FF' y también, respecto de su mediatriz. Apoyándonos en esta simetría y como A, A', B y B' son puntos de la elipse, se llega fácilmente a que:

$$a = \tfrac{1}{2}(\overline{AF} + \overline{AF'}) = \tfrac{1}{2}(\overline{AF} + \overline{FF'} + \overline{AF}) = \tfrac{1}{2}(\overline{AF} + \overline{FF'} + \overline{F'A'}) =$$
$$= \tfrac{1}{2}\overline{AA'} = \overline{OA} = \overline{OA'}$$
$$a = \tfrac{1}{2}(\overline{BF} + \overline{BF'}) = \tfrac{1}{2}(2\overline{BF}) = \overline{BF} = \overline{BF'} = \overline{B'F} = \overline{B'F'}$$

$$[OBF \text{ triángulo rectángulo}] \Rightarrow \overline{BF}^2 = \overline{OB}^2 + \overline{OF}^2 \Rightarrow a^2 = b^2 + c^2$$

Obtengamos la ecuación de la elipse. Como las coordenadas de los focos son $F(c, 0)$ y $F'(-c, 0)$, la condición $\overline{XF} + \overline{XF'} = 2a$, para que un punto $X(x, y)$ sea de la elipse, se puede poner:

$$\sqrt{(x-c)^2 + y^2} + \sqrt{(x+c)^2 + y^2} = 2a$$

Elevando al cuadrado (lo que no introduce soluciones extrañas, pues sus dos miembros son siempre positivos), se obtiene la ecuación equivalente

$$(x^2 + y^2 + c^2) + \sqrt{(x^2 + y^2 + c^2)^2 - 4c^2x^2} = 2a^2$$

Pasando $x^2 + y^2 + c^2$ al segundo miembro y elevando al cuadrado, se obtiene(*):

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2) \quad \text{o} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

que es la ecuación buscada. Finalmente, los puntos de la elipse están situados entre las dos circunferencias del enunciado ya que, para ellos, es:

$$x^2 + y^2 = a^2\left(\frac{x^2}{a^2} + \frac{y^2}{a^2}\right) \leqslant a^2\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right) = a^2$$
$$x^2 + y^2 = b^2\left(\frac{x^2}{b^2} + \frac{y^2}{b^2}\right) \geqslant b^2\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right) = b^2$$

EJERCICIO

Considérense dos ejes perpendiculares Ox y Oy y las circunferencias de centro en O y radios a y b, siendo $b < a$. Se traza una semirrecta variable que parte

(*) Existiría la posibilidad de introducir soluciones extrañas si resultase que, para alguno de los (x, y) que verifican a la ecuación de la elipse, fuese $2a^2 - (x^2 + y^2 + c^2) < 0$, o sea, si fuese $x^2 + y^2 > a^2 + b^2$; pero esto no ocurre ya que, si es $x^2/a^2 + y^2/b^2 = 1$, será

$$x^2 + y^2 = a^2\left(\frac{x^2}{a^2} + \frac{y^2}{a^2}\right) < a^2\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right) = a^2 < a^2 + b^2$$

de O, la cual corta a las circunferencias en los puntos P_a y P_b; por estos puntos se trazan perpendiculares a los ejes Ox y Oy, respectivamente, las cuales se cortan en un punto P. Hallar el lugar geométrico que describe el punto P.

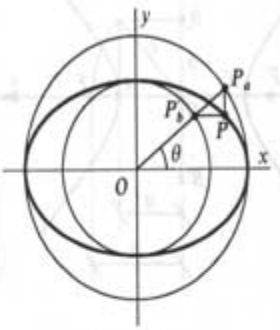


Figura 13.2.

RESOLUCION

Si llamamos θ al ángulo de Ox con la semirrecta variable, $0 \leqslant \theta < 2\pi$, las coordenadas de P_a, P_b y P serán:

$$P_a(a\cos\theta,\ a\operatorname{sen}\theta) \quad , \quad P_b(b\cos\theta,\ b\operatorname{sen}\theta) \quad , \quad P(a\cos\theta,\ b\operatorname{sen}\theta)$$

Las ecuaciones paramétricas del lugar geométrico pedido son, pues, $x = a\cos\theta$, $y = b\operatorname{sen}\theta$; eliminando el parámetro θ, se obtiene la elipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

LA HIPERBOLA; ECUACION REDUCIDA

[279]

En el plano euclídeo E_2, se llama *hipérbola* que tiene por *focos* a los puntos F y F' (situados a distancia $\overline{FF'} = 2c$) y cuya *constante* es $2a \in \mathbb{R}$ (siendo $0 < a < c$), al lugar geométrico de los puntos $X \in E_2$ tales que $|\overline{XF} - \overline{XF'}| = 2a$.

Se llaman *ejes* de la hipérbola, por ser sus ejes de simetría (ortogonal), a la recta FF' (eje focal o real) y a su mediatriz (eje secundario o imaginario); su punto de intersección, O, es su *centro* de simetría. El eje real corta a la hipérbola en dos puntos, A y A', que se llaman *vértices*; se verifica que $\overline{OA} = \overline{OA'} = a$. El eje imaginario no corta a la hipérbola, que tiene dos «ramas», una a cada lado de dicho eje.

La hipérbola, referida a sus ejes real (Ox) e imaginario (Oy), admite por ecuación a:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad (\text{donde } a^2+b^2=c^2) \text{ (ecuación reducida)} \quad \text{o}$$

$$\begin{cases} x = \pm a\ \operatorname{Ch}\theta \\ y = b\ \operatorname{Sh}\theta \end{cases}{}^{(*)} \quad \text{para } \theta \in \mathbb{R} \quad \text{(ecuaciones paramétricas)}$$

(el signo $\pm$ depende de la rama). Los puntos $B(0,\ b)$ y $B'(0,\ -b)$ se llaman «extremos del eje imaginario».

Si es $a = b$, se dice que la hipérbola es *equilátera*.

(*) El $\operatorname{Ch} x$ (coseno hiperbólico de x) y el $\operatorname{Sh} x$ (seno hiperbólico de x) se definen mediante

$$\operatorname{Ch} x = \tfrac{1}{2}(e^x + e^{-x}) \quad \text{y} \quad \operatorname{Sh} x = \tfrac{1}{2}(e^x - e^{-x}) \quad \text{(nótese que } \operatorname{Ch}^2 x - \operatorname{Sh}^2 x = 1)$$

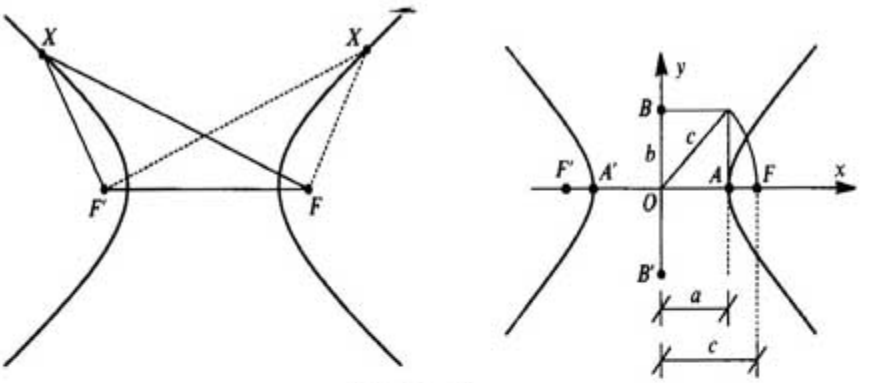


Figura 13.3.

DEMOSTRACION

La simetría de la hipérbola respecto de sus dos ejes se desprende, obviamente, de la definición. Recurriendo a dicha simetría y como A y A' son de la hipérbola, se obtiene que:

$$a = \tfrac{1}{2}(\overline{AF'} - \overline{AF}) = \tfrac{1}{2}(\overline{AA'} + \overline{A'F'} - \overline{AF}) = \tfrac{1}{2}\overline{AA'} = \overline{OA} = \overline{OA'}$$

Como los puntos del eje secundario equidistan de F y F', en él no hay puntos de la hipérbola.

Las coordenadas de los focos son $F(c, 0)$ y $F'(-c, 0)$ y, por ello, un punto $X(x, y)$ es de la hipérbola si y sólo si:

$$|\sqrt{(x-c)^2 + y^2} - \sqrt{(x+c)^2 + y^2}| = 2a$$

Elevando al cuadrado (lo que no introduce soluciones extrañas, pues sus dos miembros son positivos), se obtiene la ecuación equivalente:

$$(x^2 + y^2 + c^2) - \sqrt{(x^2 + y^2 + c^2)^2 - 4c^2x^2} = 2a^2$$

Dejando la raíz sola en el segundo miembro y elevando al cuadrado, se obtiene(*):

$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2) \quad \text{o} \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

EJERCICIO

Pruébese que, en el plano euclídeo, la elipse o la hipérbola que tiene sus focos en los puntos F y F' y cuya constante es $2a$ (para la elipse, $\overline{FF'} < 2a$; para la hipérbola $\overline{FF'} > 2a$) es el lugar geométrico de los puntos que equidistan del foco F y de la circunferencia que tiene centro en F' y radio $2a$, que se llama *circunferencia focal*. Nótese que F está en el interior o en el exterior de la circunferencia focal para la elipse o la hipérbola, respectivamente.

(*) Existiría la posibilidad de introducir alguna solución extraña si resultase que, para alguno de los (x, y) que verifican a la ecuación de la hipérbola fuese $2a^2 - (x^2 + y^2 + c^2) > 0$. Esto no es posible ya que, como $x^2 = (1 + y^2/b^2)a^2$, será

$$2a^2 - (x^2 + y^2 + c^2) = -b^2 - \frac{c^2}{a^2}y^2 < 0$$

RESOLUCION

La distancia de un punto X a la circunferencia focal de centros en F', a la que llamaremos C', es $d(X, C') = |2a - \overline{F'X}|$. Por tanto, los puntos que equidistan de F y C' son los X tales que $|2a - \overline{F'X}| = \overline{FX}$, esto es, aquellos que cumplen

$$\overline{FX} + \overline{F'X} = 2a \quad \text{o} \quad \overline{FX} - \overline{F'X} = \pm 2a \qquad [1]$$

Consideremos primero el caso de ser $\overline{FF'} < 2a$. La segunda de estas relaciones [1] no se puede verificar ahora para ningún punto X, ya que, como $\overline{FF'} \geqslant |\overline{F'X} - \overline{FX}|$ (en el triángulo $FF'X$, un lado es mayor que la diferencia de los otros dos), será $|\overline{F'X} - \overline{FX}| < 2a$ para todo X. En este caso, pues, el lugar geométrico se reduce a los puntos X tales que $\overline{FX} + \overline{F'X} = 2a$, que es la elipse.

Cuando sea $\overline{FF'} > 2a$, la primera de las relaciones [1] no se podrá verificar para ningún X, ya que, como $\overline{FX} + \overline{F'X} \geqslant \overline{FF'}$ (en el triángulo $FF'X$, un lado es menor que la suma de los otros dos), será $\overline{FX} + \overline{F'X} > 2a$ para todo X. Por ello, el lugar geométrico se reduce a los puntos X para los que se verifica $\overline{FX} - \overline{F'X} = \pm 2a$, que es la hipérbola.

ASINTOTAS DE LA HIPERBOLA

[280]

En el plano euclídeo E_2, consideremos la hipérbola que admite por ecuación reducida a:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \text{o} \quad y = \pm \frac{b}{a}\sqrt{x^2 - a^2}$$

La hipérbola tiene dos ramas (una para $x > a$ y la otra para $x < a$). Para $x \to \pm\infty$, estas ramas «tienden a confundirse»(*) con dos rectas, que se llaman *asíntotas* de la hipérbola, y que son:

$$\text{Asíntotas de la hipérbola:} \quad y = \pm \frac{b}{a} x$$

Las rectas $x = \pm a$ son (por simetría) tangntes a la hipérbola en sus vértices. Se llama *rectángulo principal* al que tiene por diagonales a las síntotas y dos de sus lados son las tangentes en los vértices; sus lados miden $2a$ y $2b$ y sus diagonales $2c$. Se dice que a y b son las longitudes del semieje real y de semieje imaginario; c es la semidistancia focal.

Se llama *hipérbola conjugada* de la dada a la que admite por ecuación a:

$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1$$

(ambas tienen los mismos ejes, las mismas asíntotas, el mismo rectángulo principal; el semieje real de una es el semieje imaginario de la otra, y recíprocamente).

(*) Para cada x se obtienen dos puntos $(x, \pm y)$ de la hipérbola; la distancia de cada uno de éstos a la correspondiente asíntota $y = (\pm b/a)x$ tiende a cero cuando $|x| \to \infty$.

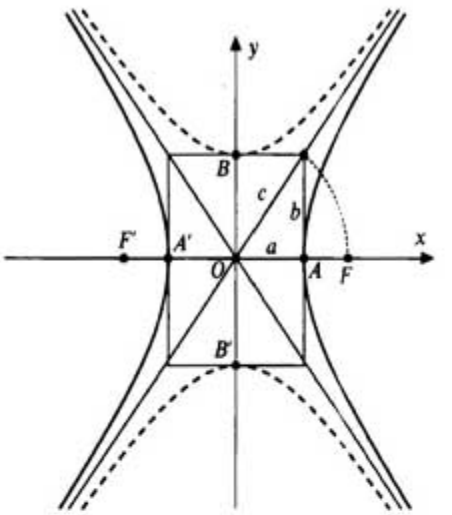


Figura 13.4.

COMPROBACION

Consideremos la asíntota correspondiente (en la indeterminación $\pm$) al signo $+$; igual se hace con la otra. En concreto, para $x \to \infty$, hemos de ver que la distancia $d(P, r)$ tiende a cero, siendo r la asíntota y P el siguiente punto de la de la hipérbola:

$$P\left(x, \frac{b}{a}\sqrt{x^2 - a^2}\right) \quad \text{y} \quad r: y = \frac{b}{a}x$$

Así ocurre, ya que

$$d(P, r) = \frac{b|\sqrt{x^2 - a^2} - x|}{\sqrt{a^2 + b^2}} = \frac{ba^2}{\sqrt{a^2 + b^2}\,|\sqrt{x^2 - a^2} + x|} =$$

$$= \frac{ba^2}{\sqrt{a^2 + b^2}} \frac{1}{|x|\,|\sqrt{1 - (a/x)^2} + 1|}$$

que evidentemente tiende a cero para $x \to \infty$.

Nótese que los lados y las diagonales del «rectángulo principal» miden lo que se dice en el enunciado debido a que las pendientes de las asíntotas son $\pm b/a$.

EJERCICIO

Hállese la intersección de una hipérbola con una recta paralela a una de las dos asíntotas de la hipérbola. Compruébese que dicha intersección consta de un solo

punto propio; véase que también tienen en común un «punto del infinito». Si la recta tiende a la asíntota, entonces el punto de intersección propio tiende también al infinito.

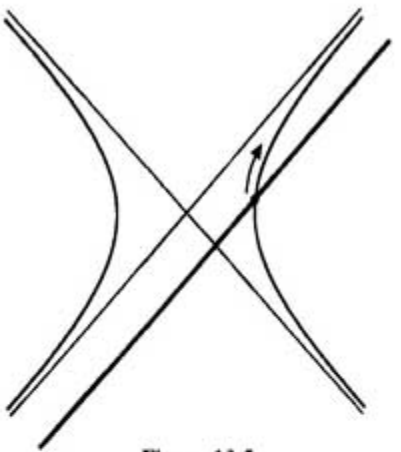


Figura 13.5.

RESOLUCION

Recurriendo a la ecuación reducida de la hipérbola, ésta y la recta del enunciado serán:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad y \quad \frac{y}{b}=\pm\frac{x}{a}+h$$

Por tanto, la intersección se obtiene de

$$\frac{x^2}{a^2}-\left(\frac{x^2}{a^2}\pm 2\,\frac{h}{a}\,x+h^2\right)=1, \quad \text{luego} \quad x=-a\,\frac{1+h^2}{2h}$$

La intersección se reduce, pues, a un solo punto propio (hay un único x para cada $h\neq 0$), salvo si la recta es la asíntota ($h=0$); si la recta tiende a la asíntota, entonces $h\to 0$ y el punto propio tiende a un punto del infinito.

Si la recta no hubiera sido paralela a la asíntota, el problema habría conducido a una ecuación de segundo grado en x (ahora no se anularía el coeficiente en x^2, como ha ocurrido antes) y habría dos soluciones para x, es decir, la intersección tendría dos puntos propios. Si esta recta (cualquiera) tiende a la recta dada (paralela a la asíntota), el coeficiente en x^2 de la anterior ecuación tendería a cero, luego una de sus raíces tendería a infinito y, por ello, uno de los dos puntos de intersección tendería a infinito, lo que nos permite decir que la paralela a la asíntota tiene en común con la hipérbola un punto propio y el otro en el infinito. A la misma conclusión se puede llegar si se utilizan coordenadas homogéneas (t, x, y); haciendo esto, la hipérbola y la recta dadas tienen por ecuaciones a:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=t^2 \quad \text{e} \quad \frac{y}{b}=\pm\frac{x}{a}+ht$$

y ambas resultan tener en común el punto $(t, x, y) = (0, a, \pm b)$, que es del infinito, pues tiene $t = 0$.

LA PARABOLA; ECUACION REDUCIDA

[281]

En el plano euclídeo E_2, se llama *parábola* que tiene por *foco* al punto F y por *directriz* a la recta d (situada a distancia $p > 0$ del foco; a p se le llama *parámetro*), al lugar geométrico de los puntos $X \in E_2$ que equidistan de F y d.

Se llama eje de la parábola, por ser su *eje* de simetría (ortogonal) a la perpendicular por F a d; el eje corta a la parábola en un punto, O, que se llama *vértice*. La parábola carece de centro. La parábola es una curva no acotada, pero no tiene asíntotas.

La parábola, referida a su eje (Ox) y a la tangente en el vértice (Oy), admite por ecuación a

$$y^2 = 2px \quad \text{(ecuación reducida)}$$

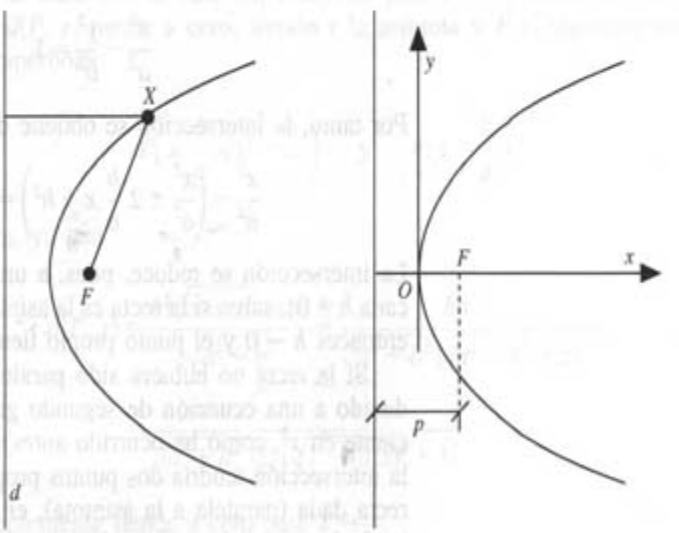


Figura 13.6.

DEMOSTRACION

El vértice O está en el eje y a distancia $p/2$ de F y, también, de d. La perpendicular al eje por O es, por razón de simetría, la tangente a la parábola en O. En los ejes Ox y Oy, las coordenadas del foco son $F(p/2, 0)$ y la ecuación

de la directriz es $d: x = -p/2$. Un punto $X(x, y)$ es, pues, de la **parábola si y** sólo si $d(X, F)$ y $d(X, d)$ son iguales, es decir:

$$\sqrt{(x-p/2)^2+y^2} = |x+p/2| \quad \text{o sea} \quad \left(x-\frac{p}{2}\right)^2 + y^2 = \left(x+\frac{p}{2}\right)^2$$

que simplificando queda $y^2 = 2px$, como había que comprobar.

EJERCICIO

En el plano euclídeo y respecto de ejes rectangulares Oxy, se consideran la elipse E y la hipérbola H que tienen a Ox como eje focal, al origen O como vértice, presentando (en él) la concavidad hacia el semieje Ox positivo, y cuyos semiejes son a y b, con $a > b$. Obtener sus ecuaciones, en los ejes Oxy, en función de $p = b^2/a$ y de $q = b^2/a^2$ (*ecuación trinomia*).

Si se varían a y b de manera que el foco correspondiente al vértice O permanezca fijo, en $F(s, 0)$, y el centro tienda a infinito (o sea, $a \to \infty$), hallar las curvas a las que «tienden» las anteriores elipse e hipérbola.

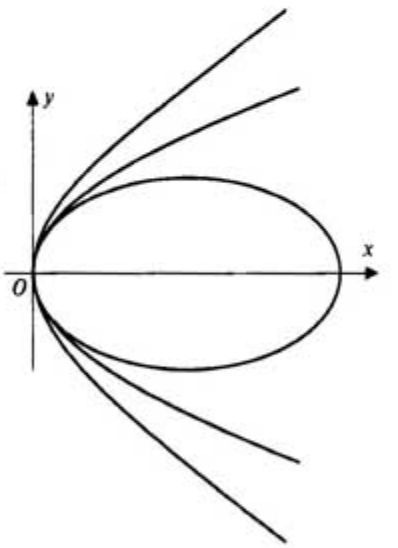


Figura 13.7.

RESOLUCION

Si en las ecuaciones reducidas de E y H (véase [278] y [279]) se traslada el origen al vértice izquierdo (sustituyendo x por $x - a$ en la elipse y por $x + a$ en la hipérbola), se obtiene

$$E: y^2 = 2\frac{b^2}{a}x - \frac{b^2}{a^2}x^2 \quad \text{o sea} \quad E: y^2 = 2px - qx^2$$

$$H: y^2 = 2\frac{b^2}{a}x + \frac{b^2}{a^2}x^2 \quad \text{o sea} \quad H: y^2 = 2px + qx^2$$

Como $s = a - c$ en la elipse y $s = c - a$ en la hipérbola, será:

- Para E:

$$p = \frac{b^2}{a} = \frac{a^2 - c^2}{a} = \frac{a^2 - (a-s)^2}{a} = 2s - \frac{s^2}{a}$$

$$q = \frac{p}{a} = \frac{2s}{a} - \frac{s^2}{a^2}$$

- Para H:

$$p = \frac{b^2}{a} = \frac{c^2 - a^2}{a} = \frac{(a+s)^2 - a^2}{a} = 2s + \frac{s^2}{a}$$

$$q = \frac{p}{a} = \frac{2s}{a} + \frac{s^2}{a^2}$$

Por tanto, al hacer que a tienda a infinito, p y q tienden (tanto para E como para H) hacia $p_\infty = 2s$ y $q_\infty = 0$. En consecuencia, la elipse y la hipérbola «tienden», ambas, a la parábola de ecuación $y^2 = 4sx$.

13.2. PRIMERAS PROPIEDADES DE LAS CONICAS

Algunas de las cuestiones que vamos a considerar aquí (la tangente a una cónica, por ejemplo) habremos de abordarlas de nuevo, cuando hagamos el estudio general de las cónicas. Para evitar excesivas reiteraciones, dichas cuestiones las trataremos ahora de manera superficial.

INTERSECCION DE UNA CONICA Y UNA RECTA

[282]

La intersección de una cónica (elipse, hipérbola o parábola) y una recta consta de dos puntos distintos (puntos propios o del infinito)(*), o de un sólo punto (doble; es decir, dos puntos «confundidos» en uno) o es vacía (dos «puntos imaginarios»). Se dice, respectivamente, que la recta es *secante*, *tangente* o *exterior* a la cónica.

(*) Si la cónica es elipse, los dos puntos son siempre propios. Si la cónica es hipérbola, uno de los dos puntos puede estar en el infinito y esto ocurre si la recta es paralela a una asíntota; si la recta es una asíntota, entonces los dos puntos están en el infinito. Si la cónica es parábola, uno de los dos puntos puede estar en el infinito; esto ocurre si la recta es paralela al eje.

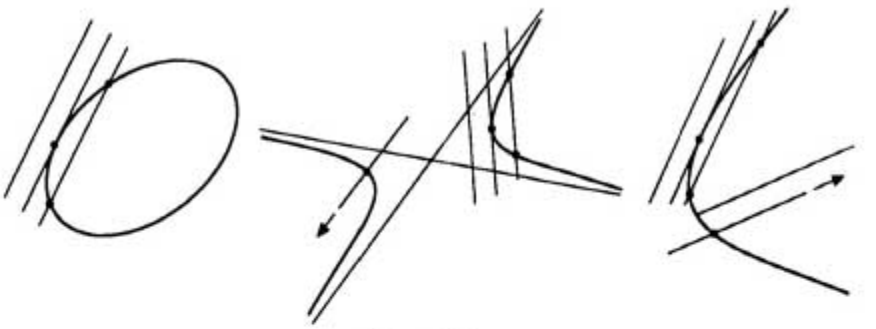


Figura 13.8.

COMPROBACION

Vamos a estudiar separadamente los casos de elipse (E), hipérbola (H) y parábola (P); tomando referencias *ad hoc*, éstas tendrán ecuaciones del tipo (reducidas):

$$E: \frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad , \quad H: \frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad , \quad P: y^2=2px \qquad [1]$$

Consideremos una recta arbitraria del plano, que definiremos por dos cualesquiera de sus puntos, a cuyas coordenadas llamaremos (x_1, y_1) y (x_2, y_2), y que admitirá ecuaciones paramétricas del tipo:

$$x=\frac{\alpha x_1+x_2}{1+\alpha} \quad , \quad y=\frac{\alpha y_1+y_2}{1+\alpha} \quad (\alpha \in \mathbb{R};\ \alpha \neq -1)^{(*)} \qquad [2]$$

Sustituyendo estos valores en las ecuaciones [1] de las cónicas, resulta que los valores de α de los puntos de intersección son las raíces de la ecuación:

$$\left.\begin{array}{l} \text{para } E: \left(\dfrac{x_1^2}{a^2}+\dfrac{y_1^2}{b^2}-1\right)\alpha^2+2\left(\dfrac{x_1x_2}{a^2}+\dfrac{y_1y_2}{b^2}-1\right)\alpha+\left(\dfrac{x_2^2}{a^2}+\dfrac{y_2^2}{b^2}-1\right)=0 \\ \text{para } H: \left(\dfrac{x_1^2}{a^2}-\dfrac{y_1^2}{b^2}-1\right)\alpha^2+2\left(\dfrac{x_1x_2}{a^2}-\dfrac{y_1y_2}{b^2}-1\right)\alpha+\left(\dfrac{x_2^2}{a^2}-\dfrac{y_2^2}{b^2}-1\right)=0 \\ \text{para } P: (y_1^2-2px_1)\alpha^2+2(y_1y_2-px_1-px_2)\alpha+(y_2^2-2px_2)=0 \end{array}\right\} \qquad [3]$$

(*) Estas paramétricas se obtienen, a partir de las paramétricas usuales $x=x_1+\lambda(x_2-x_1)$, $y=y_1+\lambda(y_2-y_1)$, haciendo el cambio de parámetro $\lambda=1:(1+\alpha)$. Nótese que, aunque éstas paramétricas no proporcionan el punto (x_1, y_1), éste se obtiene como límite para $\alpha \to \infty$. Para $\alpha \to -1$, se obtiene, como posición límite, el punto del infinito de la recta.

Estas ecuaciones [3] (en la incógnita α), como son de segundo grado, tienen dos soluciones reales distintas (recta secante), una solución real doble (recta tangente) o dos soluciones complejas conjugadas (recta exterior).

Alguno de los anteriores puntos de intersección será del infinito si alguna de las soluciones de la ecuación [3] fuese la $\alpha = -1$. En el caso de la elipse, esto no puede ocurrir pues, al poner $\alpha = -1$ en dicha ecuación, se obtendría $(x_1 - x_2)^2/a^2 + (y_1 - y_2)^2/b^2 = 0$, que no se puede verificar ya que los puntos (x_1, y_1) y (x_2, y_2) son distintos. En el caso de la hipérbola, la ecuación [3] se verifica para $\alpha = -1$ si y sólo si a $(y_2 - y_1) = \pm b(x_2 - x_1)$, esto es, si la recta dada es paralela a una de las dos asíntotas de la hipérbola. En el caso de la parábola, la ecuación [3] se verifica para $\alpha = -1$ si y sólo si $y_1 = y_2$, es decir, si la recta es paralela al eje de la parábola.

EJERCICIO

Dada una cónica (elipse, hipérbola o parábola) en el plano, considérese una recta que se mueve, paralelamente a sí misma, manteniendo siempre la misma dirección. Hallar el lugar geométrico que describe el punto medio de los dos puntos en que la recta variable corta a la cónica. Compruébese que dicho lugar es un diámetro (recta que pasa por el centro), en el caso de elipse o hipérbola, y que es una paralela al eje, en el caso de parábola.

RESOLUCION

Consideremos el caso de hipérbola (para elipse o parábola se procede de modo análogo) y expresemos ésta a través de su ecuación reducida

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

Sea m la pendiente (constante) de la recta variable, que será de la forma $y = mx + h$ (h variable); llevando esta expresión a la ecuación de la hipérbola se obtiene

$$(b^2 - a^2m^2)x^2 - 2ma^2hx - (a^2h^2 + a^2b^2) = 0$$

Si x_1 y x_2 son las dos raíces de esta ecuación, el punto medio que buscamos será el de coordenadas

$$x = \frac{x_1 + x_2}{2} \quad ; \quad y = m\,\frac{x_1 + x_2}{2} + h$$

Recurriendo a la expresión que da la suma de las raíces de una ecuación de segundo grado, se obtiene que

$$x = \frac{a^2m}{b^2 - a^2m^2}\,h \quad ; \quad y = \frac{b^2}{b^2 - a^2m^2}\,h$$

Eliminando h se obtiene el lugar pedido, que es $a^2my = b^2x$. Nótese que se trata, efectivamente, de un diámetro: el diámetro de pendiente $m' = b^2/(a^2m)$. Se dice que la dirección de pendiente m' es la *dirección conjugada* de la que tine pendiente m, respecto de la hipérbola.

TANGENTES A LAS CONICAS

[283]

En el plano euclídeo, se considera una cónica (elipse E, hipérbola H o parábola P), cuya ecuación reducida es:

$$E: \frac{x^2}{a^2}+\frac{y^2}{b^2}=1, \quad H: \frac{x^2}{a^2}-\frac{y^2}{b^2}=1, \quad P: y^2=2px$$

I. En cualquiera de sus puntos, (x_0, y_0), la cónica dada tiene tangente, que es la recta de ecuación:

$$\frac{xx_0}{a^2}+\frac{yy_0}{b^2}=1, \qquad \frac{xx_0}{a^2}-\frac{yy_0}{b^2}=1, \qquad yy_0=p(x+x_0) \qquad [A]$$

(para la elipse E) (para la hipérbola H) (para la parábola P)

II. Por un punto (x_0, y_0) que no es de la cónica, pasan dos tangentes(*) a la cónica; la ecuación conjunta de esta pareja de tangentes es:

$$\left.\begin{aligned}
&\left(\frac{xx_0}{a^2}+\frac{yy_0}{b^2}-1\right)^2=\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}-1\right)\left(\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}-1\right) && \text{(para la elipse } E\text{)}\\
&\left(\frac{xx_0}{a^2}-\frac{yy_0}{b^2}-1\right)^2=\left(\frac{x^2}{a^2}-\frac{y^2}{b^2}-1\right)\left(\frac{x_0^2}{a^2}-\frac{y_0^2}{b^2}-1\right) && \text{(para la hipérbola } H\text{)}\\
&(yy_0-p(x+x_0))^2=(y^2-2px)(y_0^2-2px_0) && \text{(para la parábola } P\text{)}
\end{aligned}\right\} \quad [B]$$

(*) Reales o imaginarias pero distintas. En el caso de la hipérbola una de las tangentes (o las dos) puede ser asíntota; de ser así, la tangencia se produce en su punto del infinito.

DEMOSTRACION

Consideraremos el caso de elipse; en los otros dos, se puede proceder de manera análoga. Tomemos una recta cualquiera, que la definiremos por dos de sus puntos (x_1, y_1) y (x_2, y_2), y procedamos como en el apartado anterior, [282], recurriendo a las ecuaciones paramétricas [2] y obteniendo la ecuación de segundo grado [3] que da los valores de α de los puntos de intersección. La recta será, pues, tangente a la elipse si las dos raíces de la ecuación [3] son iguales, esto es, si su discriminante es nulo, o sea, si

$$\left(\frac{x_1x_2}{a^2}+\frac{y_1y_2}{b^2}-1\right)^2=\left(\frac{x_1^2}{a^2}+\frac{y_1^2}{b^2}-1\right)\left(\frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}-1\right) \qquad [C]$$

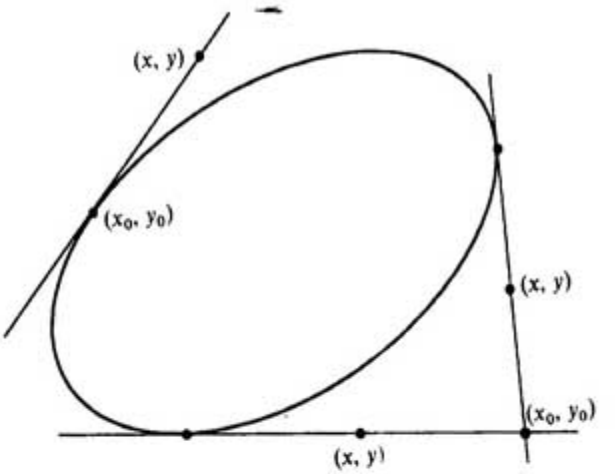


Figura 13.9.

Las tangentes a la elipse E que pasan por el punto (x_0, y_0) estarán formadas por los puntos (x, y) tales que la recta que le une con (x_0, y_0) es tangente a E, es decir, tales que se verifica [C] para (x_0, y_0) y (x, y), en lugar de (x_1, y_1) y (x_2, y_2), lo que no es otra cosa que la ecuación [B] del enunciado. Esta ecuación, que lo es de una familia de rectas, como es de segundo grado, representa a dos rectas: las dos tangentes a E desde (x_0, y_0). Si este punto (x_0, y_0) perteneciera a la elipse, entonces el último factor de la ecuación [B] sería nulo y ésta se reduciría a:

$$\left(\frac{xx_0}{a^2}+\frac{yy_0}{b^2}-1\right)^2=0$$

que es, contada dos veces, la recta [A] del enunciado.

EJERCICIO

En el plano euclídeo, se considera la elipse de ecuación reducida

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

Sea d el diámetro (recta que pasa por el centro) de pendiente m, que corta a la elipse en dos puntos; compruébese que las tangentes a la elipse en dichos puntos son paralelas y hállese su pendiente m'. Compruébese que las tangentes en los extremos del diámetro de pendiente m' tienen pendiente m. Se dice que los diámetros de pendientes m y m' son *diámetros conjugados*.

RESOLUCION

Los puntos de la elipse situados en el diámetro de pendiente m son de la forma

$P_0(\pm x_0, \pm m x_0)$. De acuerdo con la fórmula [A] de [283], las pendientes de las tangentes en los puntos citados, $\underline{P}_0$, son

$$m' = -\frac{(\pm x_0)b^2}{(\pm m x_0)a^2} = -\frac{b^2}{ma^2}$$

que son iguales (no dependen del signo $\pm$). Repitiendo lo anterior con m', en lugar de m, se obtiene un m'' que valdrá $m'' = (-b^2) : (m'a^2)$. De esta expresión y de la anterior, se deduce que $m'' = m$.

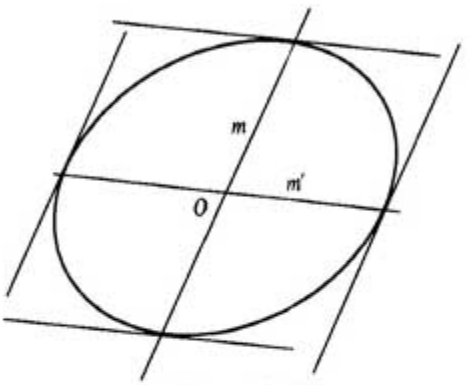


Figura 13.10.

EJERCICIO

En el plano euclídeo y usando coordenadas rectangulares xy, se considera un punto (x_0, y_0) y la cónica (una elipse E, una hipérbola H o una parábola P) cuya ecuación reducida es:

$$E: \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1; \qquad H: \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1; \qquad P: y^2 = 2px$$

La *polar* del punto (x_0, y_0) respecto de la cónica es(*) la recta que pasa por los dos puntos de tangencia, de las tangentes a la cónica desde el punto (x_0, y_0). Compruébese que dicha polar es, respectivamente, la recta:

$$\frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1 \quad ; \quad \frac{xx_0}{a^2} - \frac{yy_0}{b^2} = 1 \quad ; \quad yy_0 = p(x + x_0)$$

Nótese que, como caso límite, si (x_0, y_0) es un punto de la cónica, su tangente en (x_0, y_0) (véase [283], [A]) coincide con la polar de dicho punto.

RESOLUCION

Supongamos que la cónica es hipérbola; en los otros dos casos, se razona de igual modo. Los dos puntos de tangencia son aquellos cuyas coordenadas

(*) Este modo de introducir la polar es incompleto. Para dar una definición satisfactoria, necesitaríamos hablar de «conjugación» respecto de las cónicas, lo que nos llevaría más allá de donde queremos llegar.

verifican a las ecuaciones de H y del par de tangentes a H desde (x_0, y_0); esta última es la [B] de [283]. Por verificarse la ecuación de H, el segundo miembro de [B] es nulo y, por ello, los puntos de tangencia son las soluciones del sistema

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad [1] \qquad \text{y} \qquad \frac{xx_0}{a^2}-\frac{yy_0}{b^2}=1 \quad [2]$$

Es decir, los puntos de tangencia son los que están en la hipérbola, ecuación [1], y en la recta [B]; esta recta es, pues, la que pasa por los puntos de tangencia, es decir, la polar.

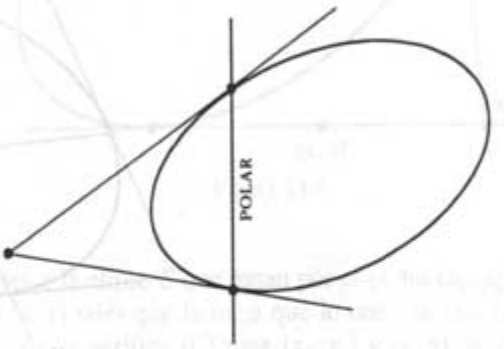


Figura 13.11.

FOCO, DIRECTRIZ Y EXCENTRICIDAD

[284]

En el plano euclídeo E_2, dados un punto F (*foco*), una recta d (*directriz*), que no pasa por F, y un número no negativo e (*excentricidad*), se verifica que el lugar geométrico de los puntos tales que sus distancias a F y d se mantienen en relación constante e igual a e es una cónica. Más exactamente:

$$\left\{X \in E_2 \Big/ \frac{d(X,\ F)}{d(X,\ d)}=e\right\}=\begin{cases}\text{elipse si } e<1\\ \text{parábola si } e=1\\ \text{hipérbola si } e>1\end{cases}$$

Para las siguientes cónicas (dadas por sus ecuaciones reducidas), los focos, las correspondientes directrices y las excentricidades son:

Cónica	Focos	Directrices	Excentricidad	
Elipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$	$F(\pm c, 0)$	$d: x=\pm\frac{a^2}{c}$	$e=\frac{c}{a}<1$	$(a^2=b^2+c^2)$
Hipérbola $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$	$F(\pm c, 0)$	$d: x=\pm\frac{a^2}{c}$	$e=\frac{c}{a}>1$	$(a^2+b^2=c^2)$
Parábola $y^2=2px$	$F(p/2, 0)$	$d: x=-p/2$	$e=1$	—

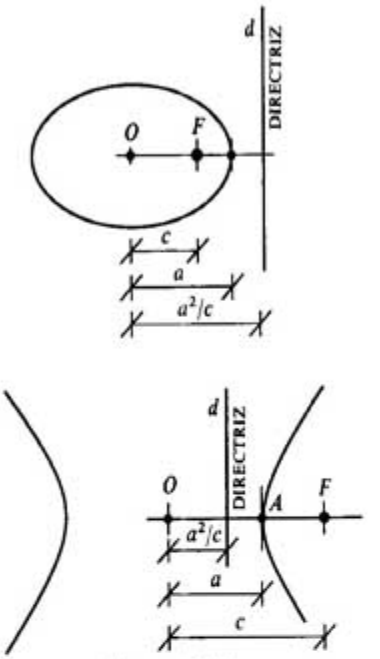


Figura 13.12.

DEMOSTRACION

Respecto de la parábola no se ha dicho ahora nada que no sepamos ya de antes (véase [281]). Las ecuaciones reducidas de la elipse y de la hipérbola pueden expresarse de las siguientes formas, equivalentes entre sí (se llama $e = c/a$):

$$x^2\frac{b^2}{a^2} \pm y^2 = b^2, \quad \pm x^2(1-e^2) \pm y^2 = b^2, \quad x^2 + y^2 = x^2e^2 \pm b^2,$$

$$(x-c)^2 + y^2 = (-2cx + c^2) + e^2x^2 \pm b^2, \quad (x-c)^2 + y^2 = e^2\left(x - \frac{a^2}{c}\right)^2,$$

$$\sqrt{(x-c)^2 + y^2} = e\left|x - \frac{a^2}{c}\right|, \quad d(X, F) = e\,d(X, d)$$

Tanto la elipse como la hipérbola es, en consecuencia, el lugar geométrico de los puntos del plano cuya relación de distancias a $F(c, 0)$ y a $d: x = a^2/c$ es constante e igual a $e = c/a$.

Recíprocamente, dados el punto F y la recta d, tomando ejes cartesianos rectangulares, xy, con origen en F y cuyo «eje de la x» sea perpendicular a d, el lugar de los puntos X tales que $d(X, F) = ed(X, d)$ resulta ser una cónica pues, como se comprueba fácilmente, la ecuación que así se obtiene es la que resulta de trasladar, en la ecuación reducida, el origen de coordenadas a un foco.

EJERCICIO

Compruébese que las tangentes a una cónica desde uno de sus focos no son reales (son imaginarias), pues resultan tener pendientes $\pm i$ (i es la unidad imaginaria). No obstante, operar formalmente y comprobar que la directriz aparece como polar (recta que une los puntos de tangencia; véase el ejercicio anterior) del correspondiente foco.

RESOLUCION

Supongamos que a la cónica se la conoce a través de su ecuación reducida y que se trata de una elipse; para hipérbola o parábola, se razona de modo análogo. Si en la ecuación conjunta de las tangentes desde un punto (x_0, y_0) (véase [289], [B]) se toma $x_0 = c$ e $y_0 = 0$ (un foco de la elipse), se obtiene la siguiente ecuación, que puede expresarse también de las formas que se indican a continuación:

$$\left(\frac{cx}{a^2} - 1\right)^2 = \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1\right)\left(\frac{-b^2}{a}\right), \quad (cx - a^2)^2 = -(b^2x^2 + a^2y^2 - a^2b^2),$$

$$a^2(x^2 - 2cx + c^2 + y^2) = 0, \quad (x - c)^2 + y^2 = 0, \quad y = \pm i(x - c)$$

Recurriendo a la ecuación de la polar que se da en el ejercicio anterior, la polar del foco $F(c, 0)$ es la recta

$$\frac{xc}{a^2} + 0 = 1, \quad \text{o sea} \quad x = \frac{a^2}{c} \quad \text{(directriz)}$$

ECUACION FOCAL DE UNA CONICA

[285]

La cónica que tiene excentricidad $e > 0$, que tiene un foco en $F(\alpha, \beta)$ y cuya directriz correspondiente es la recta $hx + ky + l = 0$ tiene por ecuación (en coordenadas rectangulares xy) a:

$$(x - \alpha)^2 + (y - \beta)^2 = e^2 \frac{(hx + ky + l)^2}{h^2 + k^2} \quad \text{(ecuación focal)}$$

COMPROBACION

La cónica está formada por los puntos $X(x, y)$ tales que

$$d(X, F) = e\, d(X, d) \quad \text{o sea} \quad \sqrt{(x - \alpha)^2 + (y - \beta)^2} = e \frac{|hx + ky + l|}{\sqrt{h^2 + k^2}}$$

Como los dos miembros de esta ecuación son positivos, elevando al cuadrado se llega a una ecuación equivalente a ella, que es la que había que obtener.

EJERCICIO

En el plano euclídeo, se considera la cónica que tiene excentricidad $e>0$, que tiene al punto F como foco y que tiene a la recta d como directriz (correspondiente al foco F); se llama parámetro de la cónica a $p=e\delta$, donde δ es la distancia del foco a la directriz. Tómense unos ejes rectangulares con origen en el foco F, con «eje de la x» la perpendicular a la directriz desde F orientada en el sentido que se aleja de la directriz y cuyo «eje de la y» es la paralela a la directriz por F orientada de manera que el giro de Fx a Fy sea positivo. Utilizando coordenadas polares (ρ, θ), esto es, llamando

$$x=\rho\cos\theta \quad \text{e} \quad y=\rho\,\text{sen}\,\theta$$

comprobar que la cónica admite por ecuación a

$$\rho=\frac{p}{1-e\cos\theta}$$

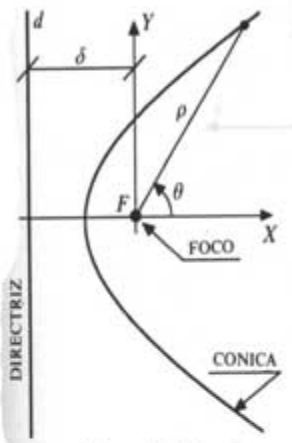


Figura 13.13.

RESOLUCION

Como en los ejes del enunciado es $F(0, 0)$ y $d: x=-\delta$, según [285] la ecuación polar de la cónica será

$$x^2+y^2=e^2(x+\delta)^2 \quad \text{o} \quad \rho^2=e^2(\rho\cos\theta+\delta)^2$$

Esta última ecuación equivale a $\rho=\pm e(\rho\cos\theta+\delta)$. El signo $\pm$ puede suprimirse ya que, si (ρ, θ) satisface a la ecuación con signo $-$, entonces $(-\rho, \theta+\pi)$ satisface a la ecuación con signo $+$ y $(-\rho, \theta+\pi)$ y (ρ, θ) representan al mismo punto. Luego la cónica admite por ecuación a

$$\rho(1-e\cos\theta)=e\delta=p \quad \text{o sea} \quad \rho=\frac{p}{1-e\cos\theta}$$

SECCIONES CONICAS

[286]

Las cónicas son las curvas de intersección de un cono de revolución con un plano que no pasa por su vértice. Si α es el semiángulo en el vértice del cono (ángulo del eje con cualquiera de las generatrices) y si β es el ángulo que forma el plano de la sección con el eje del cono, entonces la excentricidad de la cónica es $e=\cos\beta/\cos\alpha$; así pues, la cónica es elipse, parábola o hipérbola según que sea $\beta>\alpha$, $\beta=\alpha$ o $\beta<\alpha$, respectivamente.

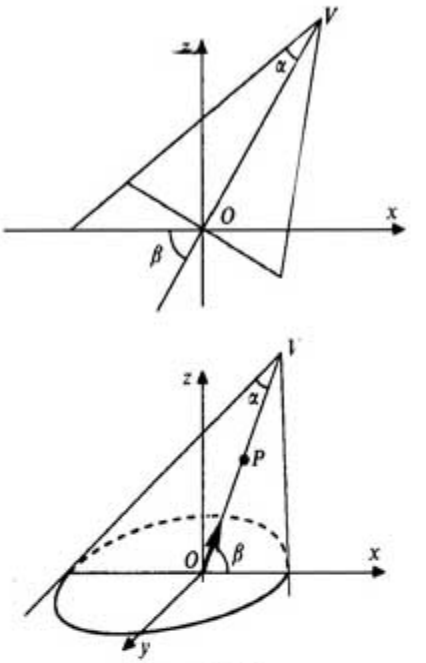


Figura 13.14.

COMPROBACION

En el espacio euclídeo E_3, tomemos unos ejes rectangulares $Oxyz$ de manera que el plano de la sección sea el $z = 0$, el origen O sea la intersección de dicho plano con el eje del cono y el vértice V del cono esté en el plano $y = 0$. Si es $h = \overline{OV}$ la «altura» del cono, las coordenadas del vértice son $V(h\cos\beta, 0, h \operatorname{sen}\beta)$. El cono está engendrado por la circunferencia, variable, de centro en un punto $P(\lambda\cos\beta, 0, \lambda \operatorname{sen}\beta)$ del eje, situada en un plano perpendicular al eje y cuyo radio es $(h - \lambda)\operatorname{tg}\alpha$. Esta circunferencia admite por ecuaciones a las:

$$\begin{cases}(x - \lambda\cos\beta)^2 + y^2 + (z - \lambda \operatorname{sen}\beta)^2 = (h - \lambda)^2 \operatorname{tg}^2\alpha \\ \cos\beta(x - \lambda\cos\beta) + \operatorname{sen}\beta(z - \lambda \operatorname{sen}\beta) = 0\end{cases}$$

Eliminando λ entre estas dos ecuaciones, se obtiene la ecuación del cono, que resulta ser

$$x^2 + y^2 + z^2 - (x\cos\beta + z \operatorname{sen}\beta)^2 = (h - x\cos\beta - y \operatorname{sen}\beta)^2 \operatorname{tg}^2\alpha$$

La sección del cono por el plano dado (el $z = 0$) será, pues:

$$\begin{cases}x^2(1 - e^2) + y^2 = -(2h\cos\beta \operatorname{tg}^2\alpha)x + h^2 \operatorname{tg}^2\alpha \quad (e = \cos\beta/\cos\alpha) \\ z = 0\end{cases}$$

que es una cónica que tiene excentricidad e. Dejamos como ejercicio la comprobación de que, dada una cónica cualquiera, existen ciertos α, β y h tales que las anteriores ecuaciones representa a la cónica dada (en una referencia adecuada).

β ESTUDIO PARTICULAR DE LAS CUADRICAS

Aquí, en esta introducción a las cuádricas, vamos a abordar menos cuestiones de las que, en el caso de las cónicas, nos ocupamos en el epígrafe anterior. Hay, para ello, un motivo importante: la dificultad, de los temas tratados para las cónicas, aumenta notablemente cuando se estudian para cuádricas; a ello contribuye el hecho de que, a estas alturas, nuestros conocimientos sobre las cónicas son escasos. Así, por ejemplo, comprobar que, en general, la intersección de una cuádrica con un plano es una cónica y, más aún, analizar si ésta es elipse, parábola o hipérbola, nos iba a costar ahora muchos más esfuerzos de lo razonable. Aplazamos, pues, éstos y otros asuntos para más adelante, donde su análisis ha de resultarnos bastante más fácil.

13.3. LAS CINCO CUADRICAS

❑ EL ELIPSOIDE; ECUACION REDUCIDA

[287]

En el espacio euclídeo E_3 y respecto de ejes rectangulares $Oxyz$, se consideran los puntos $A(a, 0, 0)$, $A'(-a, 0, 0)$, $B(0, b, 0)$, $B'(0, -b, 0)$, $C(0, 0, c)$ y $C'(0, 0, -c)$, donde $a>0$, $b>0$ y $c>0$ son dados. Sean E_x, E_y, E_z las elipses cuyos ejes son BB' y CC', CC' y AA', AA' y BB', respectivamente. Se llama *elipsoide* que tiene por vértices a A, A', B, B' C y C' al lugar geométrico definido, entre otras, de las siguientes maneras, equivalentes entre sí:

1. Lugar geométrico que describe una elipse, variable, situada en un plano perpendicular el eje Oz y que tiene sus vértices en las elipses E_x y E_y.
2. Lugar geométrico que describe una elipse, variable, situada en un plano que contiene al eje Oz, que tiene dos vértices en C y C' y los otros dos en la elipse E_z.

Dicho elipsoide, referido a los ejes rectangulares $Oxyz$, admite por ecuación a:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1 \quad \text{(ecuación reducida)}$$

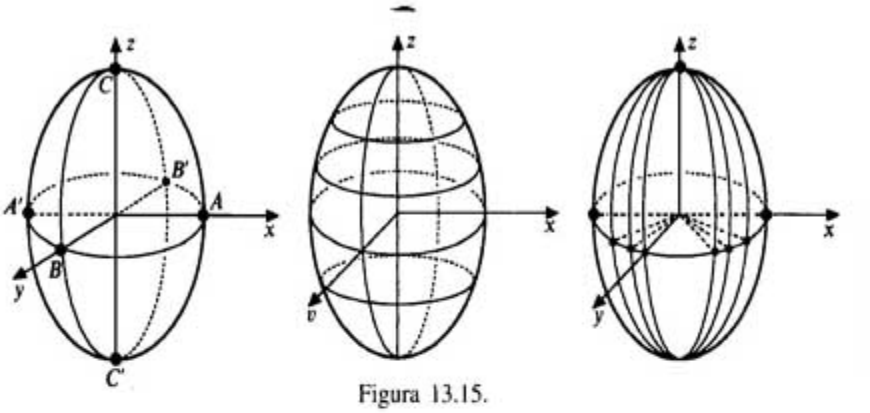


Figura 13.15.

COMPROBACION

Antes de nada, es conveniente observar que, en las anteriores definiciones 1 y 2, se pueden intercambiar los papeles que desempeñan las variables x, y, z, obteniéndose con ello el mismo elipsoide.

Para comprobar que las dos definiciones, 1 y 2, son equivalentes, vamos a obtener las ecuaciones de los dos lugares geométricos, los que se definen con una y otra; veremos que estas ecuaciones son, ambas, la que en el enunciado se llama «ecuación reducida».

Las elipses E_x, E_y y E_z admiten por ecuaciones a las:

$$E_x: \begin{cases} \dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1 \\ x=0 \end{cases} \quad ; \quad E_y: \begin{cases} \dfrac{x^2}{a^2}+\dfrac{z^2}{c^2}=1 \\ y=0 \end{cases} \quad ; \quad E_z: \begin{cases} \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1 \\ z=0 \end{cases}$$

1. Un plano perpendicular al eje Oz, el $z=h$. corta a las elipses E_x y E_y en los puntos

$$(0,\ \pm b\sqrt{1-(h/c)^2},\ h) \quad \text{y} \quad (\pm a\sqrt{1-(h/c)^2},\ 0,\ h)$$

respectivamente. Por ello la elipse variable será la:

$$\left.\begin{array}{l} \dfrac{x^2}{a^2\left(1-\dfrac{h^2}{c^2}\right)}+\dfrac{y^2}{b^2\left(1-\dfrac{h^2}{c^2}\right)}=1 \\ z=h \end{array}\right\} \qquad [1]$$

Eliminando h entre las dos ecuaciones [1], se obtiene la ecuación del lugar geométrico; ésta es, obviamente, la ecuación reducida, que había que obtener.

2. Sea θ el ángulo que forma el plano variable, π, con el plano $y=0$; tomando en el plano π ejes rectangulares Otz, donde Ot es su intersección

con el plano $z = 0$, para los puntos del plano variable π es $x = t \cos \theta$ e $y = t \operatorname{sen} \theta$. Este plano π corta a la elipse E_z en los puntos $z = 0$, $t = \pm t_0$, siendo

$$\frac{(t_0 \cos \theta)^2}{a^2} + \frac{(t_0 \operatorname{sen} \theta)^2}{b^2} = 1, \quad \text{luego} \quad t_0^2 = \frac{1}{(\cos \theta/a)^2 + (\operatorname{sen} \theta/b)^2}$$

Por tanto, en los ejes Otz, la elipse variable admite por ecuación a:

$$\frac{t^2}{t_0^2} + \frac{z^2}{c^2} = 1 \quad \text{o sea} \quad \frac{t^2 \cos^2 \theta}{a^2} + \frac{t^2 \operatorname{sen}^2 \theta}{b^2} + \frac{z^2}{c^2} = 1$$

Eliminando t y θ entre esta última ecuación y las relaciones $x = t \cos \theta$ e $y = t \operatorname{sen} \theta$, se obtiene la ecuación del lugar geométrico, que coincide con el anterior (el obtenido en 1).

[288]

El elipsoide que, respecto de unos ciertos ejes rectangulares $Oxyz$, admite por ecuación reducida a

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

es simétrico respecto del origen O (que es su *centro*), respecto de Ox de Oy y de Oz (que son sus *ejes*) y de $x = 0$, de $y = 0$ y de $z = 0$ (*planos principales*). Las secciones del elipsoide por planos secantes paralelos al $z = 0$ (o al $x = 0$ o al $y = 0$) son elipses homotéticas entre sí(*).

El elipsoide está situado entre las esferas de su mismo centro y cuyos radios son el mayor y el menor de los números, a, b y c.

Si es $a = b$, el elipsoide es de revolución (alrededor de Oz). Si $a = b = c$, el elipsoide es una esfera.

(*) Más adelante se probará que un plano secante con un elipsoide le corta siempre según una elipse y que las elipses que se obtienen al cortarle por planos paralelos son homotéticas.

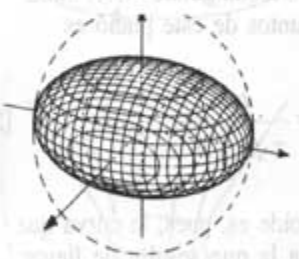


Figura 13.16.

COMPROBACION

Las simetrías del elipsoide se comprueban fácilmente sin más que observar que, si (x, y, z) es un tanto del elipsoide, también lo son todos los $(\pm x, \pm y, \pm z)$, pues satisfacen también a su ecuación.

Al cortar el elipsoide por un plano $z = h$, se obtiene la elipse que admite las anteriores ecuaciones [1], de la pág. 574, cuyos semiejes son proporcionales a a y b (la constante de proporcionalidad es $\sqrt{1 - (h/c)^2}$); estas elipses son, pues, todas homotéticas a la que se obtiene para $z = 0$.

Supongamos $a > b > c$; de ser así, para todo punto del elipsoide es:

$$x^2 + y^2 + z^2 = a^2\left(\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{a^2}\right) \leqslant a^2\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) = a^2$$

$$(x^2 + y^2 + z^2) = c^2\left(\frac{x^2}{c^2} + \frac{y^2}{c^2} + \frac{z^2}{c^2}\right) \geqslant c^2\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) = c^2$$

luego el elipsoide está situado entre las esferas de centro O y radios a y c.

EJERCICIO

Considérese el elipsoide que, respecto de ejes rectangulares $Oxyz$ del espacio euclídeo, tiene por ecuación reducida a

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad \text{(siendo } a < b < c\text{)} \qquad [1]$$

Hállese un plano que contenga al eje Oy y que corte al elipsoide según una circunferencia (a esta sección del elipsoide y a todas sus secciones paralelas se las llama *secciones cíclicas*).

RESOLUCION

Sea $z = my$ el plano buscado y tomemos en él ejes rectangulares Oty, donde Ot es su intersección con el plano $y = 0$; para los puntos de este plano es

$$x = \frac{t}{\pm\sqrt{1 + m^2}}, \quad y = y, \quad z = \frac{tm}{\pm\sqrt{1 + m^2}} \qquad [2]$$

La intersección del plano $z = my$ con el elipsoide es, pues, la curva que en los ejes rectangulares Oty admite por ecuación a la que resulta de llevar las expresiones [2] a la ecuación [1], esto es:

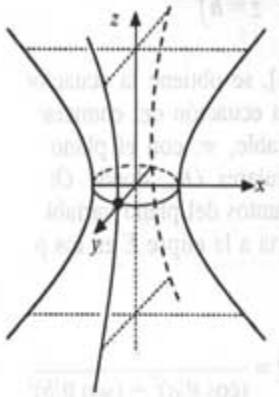
$$t^2\left[\frac{1}{(1 + m^2)a^2} + \frac{m^2}{(1 + m^2)c^2}\right] + y^2\frac{1}{b^2} = 1$$

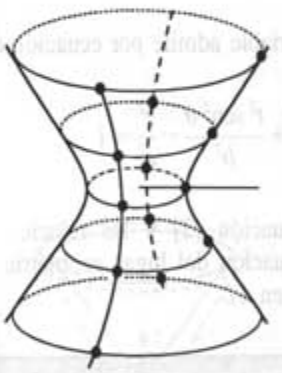
que es una elipse. Para obligar a que, según se nos pide, esta elipse sea una circunferencia, basta con igualar los coeficientes de t^2 e y^2:

$$\frac{c^2 + a^2m^2}{(1 + m^2)a^2c^2} = \frac{1}{b^2}, \quad \text{luego} \quad m = \frac{c}{a}\sqrt{\frac{b^2 - a^2}{c^2 - b^2}}$$

HIPERBOLOIDE DE UNA HOJA; ECUACION REDUCIDA

[289]





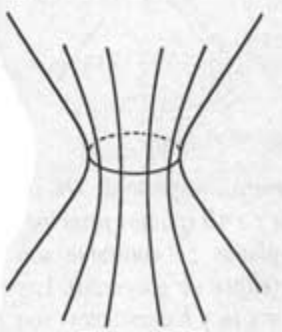


Figura 13.17.

En el espacio euclídeo E_3 y respecto de ejes rectangulares $Oxyz$, se consideran los puntos $A(a, 0, 0)$, $A'(-a, 0, 0)$, $B(0, b, 0)$, $B'(0, -b, 0)$, $C(0, 0, c)$ y $C'(0, 0, -c)$, donde $a>0$, $b>0$ y $c>0$ son dados. Sea E la elipse cuyos semiejes son AA' y BB' y H_x y H_y las hipérbolas cuyos ejes reales son BB' y AA', respectivamente, y cuyo eje imaginario es CC', para ambas. Se llame *hiperboloide de una hoja*(*) que tiene por ejes reales a AA' y BB' y por eje imaginario a CC' al lugar geométrico definido, entre otras, de las siguientes maneras, equivalentes entre sí:

1. Lugar geométrico que describe una elipse, variable, situada en un plano perpendicular al eje Oz y que tiene sus vértices en las hipérbolas H_x y H_y.
2. Lugar geométrico que describe una hipérbola, variable, situada en un plano que contiene al eje Oz, que tiene sus vértices en E y sus extremos del eje imaginario son C y C'.

El hiperboloide de una hoja, referido a los ejes $Oxyz$, admite por ecuación a:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad \text{(ecuación reducida)}$$

(*) Más adelante se verá que este hiperboloide también debe ser llamado «hiperboloide reglado» e «hiperboloide hiperbólico».

COMPROBACION

Nótese que, en la definición de hiperboloide de una hoja que acaba de darse, son intercambiables las variables y y z pero que la x desempeña un papel esencialmente distinto del de aquéllas.

Para comprobar que las dos definiciones anteriores, 1 y 2, son equivalentes, obtendremos las ecuaciones de los dos lugares geométricos, los que se definen con una y otra, y veremos que son la misma, que coinciden con la que en el enunciado se llamó «ecuación reducida».

La elipse E y las hipérbolas H_x y H_y admiten por ecuaciones a las:

$$E:\begin{cases}\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1\\ z=0\end{cases} \quad ; \quad H_x:\begin{cases}\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1\\ x=0\end{cases} \quad ; \quad H_y:\begin{cases}\dfrac{x^2}{a^2}-\dfrac{z^2}{c^2}=1\\ y=0\end{cases}$$

1. Un plano perpendicular al eje Oz, el $z=h$, corta a las hipérbolas H_x y H_y en los puntos

$$(0, \pm b\sqrt{1+(h/c)^2}, h) \quad \text{y} \quad (\pm a\sqrt{1+(h/c)^2}, 0, h)$$

respectivamente. Por ello la elipse variable será la:

$$\left.\begin{array}{l} \dfrac{x^2}{a^2\left(1+\dfrac{h^2}{c^2}\right)}+\dfrac{y^2}{b^2\left(1+\dfrac{h^2}{c^2}\right)}=1 \\ z=h \end{array}\right\} \qquad [1]$$

Eliminando h entre las dos ecuaciones [1], se obtiene la ecuación del lugar geométrico; al hacerlo, se obtiene la ecuación del enunciado.

2. Sea θ el ángulo que forma el plano variable, π, con el plano $y=0$; tomando en dicho plano π ejes rectangulares Otz, donde Ot es su intersección con el plano $z=0$, para los puntos del plano variable π es $x=t\cos\theta$ e $y=t\,\text{sen}\,\theta$. Este plano π corta a la elipse E en los puntos $z=0$, $t=\pm t_0$, siendo

$$\frac{(t_0\cos\theta)^2}{a^2}+\frac{(t_0\,\text{sen}\,\theta)^2}{b^2}=1, \quad \text{luego} \quad t_0^2=\frac{1}{(\cos\theta/a)^2+(\text{sen}\,\theta/b)^2}$$

Por tanto, en los ejes Otz, la hipérbola variable admite por ecuación a:

$$\frac{t^2}{t_0^2}+\frac{z^2}{c^2}=1 \quad \text{o sea} \quad \frac{t^2\cos^2\theta}{a^2}+\frac{t^2\,\text{sen}^2\,\theta}{b^2}-\frac{z^2}{c^2}=1 \qquad [2]$$

Eliminando t y θ entre esta última ecuación [2] y las relaciones $x=t\cos\theta$ e $y=t\,\text{sen}\,\theta$, se obtiene la ecuación del lugar geométrico, que coincide con el anterior (el obtenido en 1).

[290]

El hiperboloide de una hoja que, respecto de unos ciertos ejes rectangulares $Oxyz$, admite por ecuación reducida a

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

es simétrico respecto del origen O (que es su *centro*), respecto de Ox, de Oy y de Oz (que son sus *ejes*) y de $x=0$, $y=0$ y $z=0$ (*planos principales*). Las secciones del hiperboloide por los planos $z=$ constante son elipses homotéticas a la elipse del plano $z=0$ (*elipse de garganta*). Las secciones del hiperboloide por planos $x=h$ e $y=k$ (h y k constantes) son hipérbolas homotéticas, respectivamente, a las de los planos $x=0$ e $y=0$, si $|h|<a$ y $|k|<b$, o a sus hipérbolas conjugadas, si $|h|>a$ y $|k|>b$(*).

Si es $a=b$, el hiperboloide es de revolución (alrededor del eje Oz).

(*) Si $h=\pm a$ o $k=\pm b$, la intersección es un par de rectas; más adelante se analizará este hecho.

COMPROBACION

Para verificar cuanto se acaba de decir, se pueden seguir los mismos pasos que se dieron en la comprobación de las correspondientes propiedades de los elipsoides (véase [288]). Para evitar reiteraciones, nos limitaremos a comprobar lo referente a los cortes del hiperboloide por planos $x = h$ (para los planos $y = k$, se razona de modo análogo), que es donde se presenta alguna novedad.

La intersección del hiperboloide por el plano $x = h$ es la curva de ecuaciones:

$$\left.\begin{aligned} \frac{y^2}{b^2[1-(h/a)^2]} - \frac{z^2}{c^2[1-(h/a)^2]} = 1 \\ x = h \end{aligned}\right\}$$

que una hipérbola cuyos semiejes real e imaginario son

$$\begin{cases} b\sqrt{1-(h/a)^2} & \text{y} \quad c\sqrt{1-(h/a)^2}, & \text{si} \quad |h| < a \\ c\sqrt{(h/a)^2-1} & \text{y} \quad b\sqrt{(h/a)^2-1}, & \text{si} \quad |h| > a \end{cases}$$

luego esta hipérbola es homotética a la correspondiente a $x = 0$ (cuyos semiejes real e imaginario son b y c) si es $|h| < a$ y es homotética a la hipérbola conjugada de ésta (cuyos semiejes real e imaginario son c y b) si es $|h| > a$.

EJERCICIO

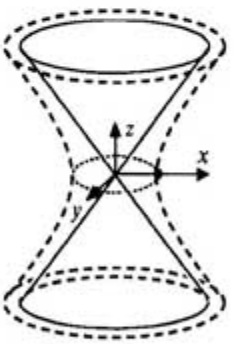


Figura 13.18.

Considérese el hiperboloide de una hoja que, respecto de ejes rectangulares $Oxyz$ del espacio euclídeo, tiene por ecuación reducida a la:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

Hállese el lugar geométrico que engendran las asíntotas de las hipérbolas que resultan de cortar el hiperboloide por los planos que contienen al eje Oz. Compruébese que dicho lugar es un cono (que se llama *cono asintótico* del hiperboloide).

RESOLUCION

Llamemos θ al ángulo que forma el plano variable π, que se cita en el enunciado, con el plano $y = 0$ (θ varía de 0 a 2π); tomemos en el plano π ejes rectangulares Otz, donde Ot es su intersección con el plano $z = 0$. Para los puntos de π es $x = t\cos\theta$ e $y = t\,\text{sen}\,\theta$ y, por tanto, en los ejes Otz la ecuación de la intersección es:

$$t^2\left(\frac{\cos^2\theta}{a^2} + \frac{\text{sen}^2\,\theta}{b^2}\right) - \frac{z^2}{c^2} = 1$$

Las asíntotas de esta hipérbola son las rectas que admiten por ecuación, conjunta, a

$$t^2\left(\frac{\cos^2\theta}{a^2}+\frac{\operatorname{sen}^2\theta}{b^2}\right)-\frac{z^2}{c^2}=0$$

Por tanto, la ecuación del lugar geométrico que engendran estas rectas se obtiene eliminando t y θ entre la última ecuación y las $x=t\cos\theta$ e $y=t\operatorname{sen}\theta$. Al eliminar dichos parámetros se obtiene:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0 \qquad [1]$$

Obsérvese que esta última ecuación lo es del cono que, desde el origen de coordenadas, proyecta a la elipse que admite por ecuaciones a las:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad \text{y} \quad z=c$$

Para comprobar este aserto, téngase en cuenta que este cono está formado por los puntos $x=\alpha\lambda$, $y=\beta\lambda$, $z=\gamma\lambda$, donde $\lambda\in\mathbb{R}$ y (α, β, γ) es un punto de dicho elipse, es decir tal que $(\alpha^2/a^2)+(\beta^2/b^2)=1$ y $\gamma=c$; eliminando λ, α, β y γ entre estas últimas (cinco) ecuaciones, se obtiene la ecuación [1].

HIPERBOLOIDE DE DOS HOJAS; ECUACION REDUCIDA

[291]

En el espacio euclídeo E_3 y respecto de ejes rectangulares $Oxyz$, se consideran los puntos $A(a, 0, 0)$, $A'(-a, 0, 0)$, $B(0, b, 0)$, $B'(0,-b, 0)$, $C(0, 0, c)$ y $C'(0, 0, -c)$, donde $a>0$, $b>0$ y $c>0$ son dados. Sean H_y y H_z las hipérbolas cuyo eje real es AA', para ambas, y cuyos ejes imaginarios son BB' y CC', respectivamente. Se llama *hiperboloide de dos hojas*(*), que tiene por eje real a AA' y por ejes imaginarios a BB' y CC', al lugar geométrico que describe una elipse, variable, situada en un plano perpendicular al eje Oyz que tiene sus vértices en las hipérbolas H_y y H_z.

El hiperboloide de dos hojas, referido a los ejes $Oxyz$, admite por ecuación a

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad \text{(ecuación reducida)}$$

(*) Más adelante se verá que este hiperboloide también debe ser llamado «hiperboloide elíptico» e «hiperboloide no reglado».

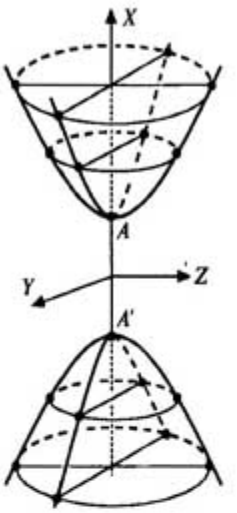


Figura 13.19.

COMPROBACION

Para probar que la ecuación del hiperboloide de dos hojas es la que, en el enunciado, se llama «ecuación reducida», basta con repetir, con mínimos retoques, lo que ya se dijo para hacer, en [289]. 1, la correspondiente comprobación en el caso de hiperboloide de una hoja.

[292]

El hiperboloide de dos hojas que, respecto de unos ciertos ejes rectangulares *Oxyz*, admite por ecuación reducida a

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

es simétrico respecto del origen *O* (que es su *centro*), respecto de *Ox*, de *Oy* y de *Oz* (que son sus *ejes*) y de $x=0$, $y=0$ y $z=0$ (*planos principales*).

Las secciones del hiperboloide por los planos $x=h$ (h constante y tal que $|h|>a$) son elipses homotéticas entre sí(*). Las secciones del hiperboloide por planos perpendiculares a *Oy* y a *Oz* son hipérbolas homotéticas a las de los planos $y=0$ y $z=0$, respectivamente.

Si es $b=c$, el hiperboloide es de revolución (alrededor de *Ox*).

(*) La intersección del hiperboloide con $x=a$ o con $x=-a$ sólo tiene un punto real (es un par de rectas imaginarias conjugadas con intersección real); sobre esto se hablará más adelante.

COMPROBACION

Lo que se acaba de decir puede comprobarse como en [288], sin más que introducir leves retoques.

EJERCICIO

Considérese el hiperboloide de dos hojas que, respecto de ejes rectangulares $Oxyz$, tiene por ecuación reducida a la:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

1. Hallar el lugar geométrico que engendran las asíntotas a las hipérbolas que resultan de cortar el hiperboloide por los planos que contienen al eje Ox. Compruébese que dicho lugar es un cono (que se llama cono asintótico del hiperboloide.
2. Hallar un plano que pase por el vértice $A(a, 0, 0)$, que sea paralelo al eje Oy y que corte al hiperboloide según una parábola.

RESOLUCION

1. Procediendo como se hizo en el ejercicio anterior (el de [290]), se obtiene, como allí, que el lugar geométrico pedido es el cono de ecuación:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=0$$

2. Sea θ el ángulo que forma el plano buscado, π, con el eje Oy; tomemos en el plano π unos ejes rectangulares Ayt, donde Ay es paralelo a Oy y At es la intersección de π con $y=0$. Para los puntos de π es $x=a+t\cos\theta$ y $z=t\operatorname{sen}\theta$. Por tanto, en los ejes Ayt de su plano, la ecuación de la intersección es:

$$\frac{(a+t\cos\theta)^2}{a^2}-\frac{y^2}{b^2}-\frac{(t\operatorname{sen}\theta)^2}{c^2}=1$$

que es una cónica y puede ponerse en la forma

$$y^2\left(\frac{1}{b^2}\right)=2\,\frac{\cos\theta}{a}\,t+t^2\left(\frac{\cos^2\theta}{a^2}-\frac{\operatorname{sen}^2\theta}{c^2}\right)$$

Esta cónica será parábola si se anula el término en t^2, esto es, si $\cos^2\theta/a^2=\operatorname{sen}^2\theta/c^2$, o sea, si $\operatorname{tg}\theta=\pm c/a$. Por tanto, hay dos soluciones al problema, que son los planos

$$z=\pm\,\frac{c}{a}\,(x-1)$$

PARABOLOIDE ELIPTICO

[293]

En el espacio euclídeo E_3 y respecto de ejes rectangulares $Oxyz$, se consideran las parábolas P_1 y P_2 cuyas ecuaciones reducidas son

$$P_1 \begin{cases} x^2 = 2pz \\ y = 0 \end{cases} \quad \text{y} \quad P_2 \begin{cases} y^2 = 2qz \\ x = 0 \end{cases} \quad (p>0;\ q>0)$$

(tienen el mismo vértice y el mismo semieje, con la concavidad hacia el mismo sentido; están situadas en planos perpendiculares). Se llama *paraboloide elíptico*(*), definido por las parábolas dadas, al lugar geométrico que describe una prábola P, variable, paralela a P_2, que se desplaza de manera que su vértice recorre P_1. Este paraboloide admite por ecuación a:

$$2z = \frac{x^2}{p} + \frac{y^2}{q} \quad \text{(ecuación reducida)}$$

El paraboloide elíptico, que carece de centro, es simétrico respecto de Oz (que es su *eje*) y respecto de los planos $x = 0$ e $y = 0$ (*planos principales*). Las secciones del paraboloide por planos $z = h$ (h contante, $h>0$) son elipses homotéticas entre sí; sus secciones por planos paralelos a Oz son parábolas.

Si $p = q$, el paraboloide es de revolución (alrededor de Oz).

(*) Más adelante se verá que también se le deberá llamar «paraboloide no reglado».

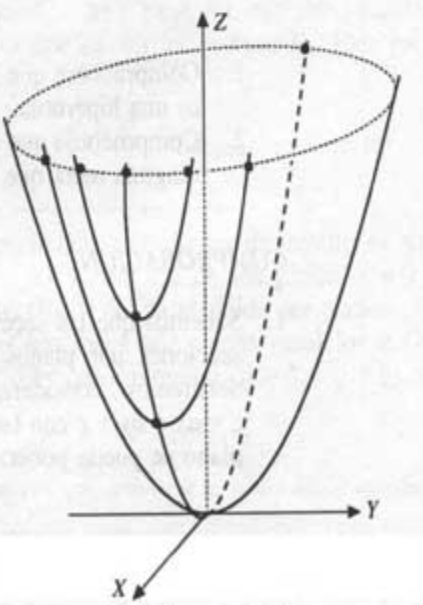


Figura 13.20.

COMPROBACION

Un punto X será del paraboloide si se puede poner $\overrightarrow{OX} = \overrightarrow{OX_1} + \overrightarrow{OX_2}$, donde $X_1 \in P_1$ y $X_2 \in P_2$. Recurriendo a dos parámetros λ, $\mu \in \mathbb{R}$, se puede poner:

$$\left.\begin{matrix}\overrightarrow{OX_1}(\lambda,\ 0,\ \lambda^2/2p)\\ \overrightarrow{OX_2}(0,\ \mu,\ \mu^2/2q)\end{matrix}\right\} \quad \text{luego} \quad X\begin{cases}x = \lambda & +0\\ y = 0 & +\mu\\ z = \lambda^2/2p + \mu^2/2q\end{cases}$$

Al eliminar λ y μ entre estas tres ecuaciones, se obtiene la ecuación reducida del enunciado.

Las simetrías del paraboloide son evidentes. Respecto de sus secciones por planos perpendiculares y paralelos a Oz, como éstas son, respectivamente:

$$\left.\begin{matrix}\dfrac{x^2}{2hp} + \dfrac{y^2}{2hq} = 1\\ z = h \quad (h > 0)\end{matrix}\right\} \qquad \left.\begin{matrix}x^2 = 2\,\dfrac{pq}{q + pm^2}\,z\\ y = mx\end{matrix}\right\}$$

resulta evidente que son, respectivamente, elipses (homotéticas al variar h) y parábolas.

EJERCICIO

Considérese el paraboloide elíptico que, en ciertos ejes $Oxyz$ rectangulares, admite por ecuación reducida a

$$2z = \frac{x^2}{p} + \frac{y^2}{q} \quad (\text{donde } p > 0 \text{ y } q > 0)$$

1. Compruébese que ninguna de las secciones planas de dicho paraboloide es una hipérbola.
2. Compruébese que este paraboloide no es reglado, es decir, que no existe ninguna recta que esté incluida en él.

COMPROBACION

1. Sabemos que las secciones por planos $z =$ constante son elipses y que las secciones por planos paralelos al eje Oz son parábolas. Por ello, sólo deberemos considerar secciones por otros planos, que serán los $z = ax + by + c$ con $(a, b) \neq (0, 0)$. La sección S del paraboloide por un tal plano se puede poner:

$$S:\begin{cases}\dfrac{x^2}{p} + \dfrac{y^2}{q} = 2ax + 2by + 2c\\ z = ax + by + c\end{cases}$$

La proyección de S sobre el plano $z=0$ es la curva

$$S_z:\begin{cases}\dfrac{x^2}{p}+\dfrac{y^2}{q}=2ax+2by+2c\\ z=0\end{cases}$$

Como esta curva es una elipse (por ser $p>0$ y $q>0$; su ecuación reducida se obtiene trasladando los ejes), la curva S no puede ser hipérbola(*) pues al proyectarla no tiene ramas infinitas.

2. Supongamos que hubiera una recta r incluida en el paraboloide. Esta recta no puede ser paralela al plano $z=0$ (pues las secciones del paraboloide por planos $z=$ constante son elipses), luego r corta al plano $z=0$. Como en este plano el paraboloide sólo tiene el punto (0, 0, 0), la recta r deberá pasar por él; es decir, r tendrá ecuaciones del tipo $x=\lambda u$, $y=\lambda v$, $z=\lambda w$ (donde u, v y w son fijos y $\lambda\in\mathbb{R}$). Al hallar la intersección de esta recta con el paraboloide se obtiene (en todos los casos) sólo dos puntos (correspondientes a dos valores de λ), como se comprueba trivialmente, luego r no puede estar incluida en el paraboloide.

PARABOLOIDE HIPERBOLICO

[294]

En el espacio euclídeo E_3 y respecto de ejes rectangulares $Oxyz$, se consideran las parábolas P_1 y P_2 cuyas ecuaciones reducidas son:

$$P_1:\begin{cases}x^2=2pz\\ y=0\end{cases}\quad\text{y}\quad P_2:\begin{cases}y^2=-2qz\\ x=0\end{cases}\qquad(p>0;\ q>0)$$

(tiene el mismo vértice; sus semiejes son opuestos, tienen la concavidad en distinto sentido; están situadas en planos perpendiculares). Se llama *paraboloide hiperbólico*(*), definido por las parábolas dadas, al lugar geométrico que describe una parábola, variable, paralela a P_2, que se desplaza de manera que su vértice recorre P_1. Este paraboloide admite por ecuación a:

$$2z=\frac{x^2}{p}-\frac{y^2}{q}\qquad\text{(ecuación reducida)}$$

El paraboloide hiperbólico, que carece de centro, es simétrico respecto de Oz (que es su *eje*) y respecto de los planos $x=0$ e $y=0$ (*planos principales*). Las secciones del paraboloide por planos $z=h$ (h constante) son hipérbolas; sus secciones por planos paralelos a Oz son parábolas. Este paraboloide no es (para ningunos $p>0$ y $q>0$) una superficie de revolución.

(*) Más adelante se verá que también se le deberá llamar «paraboloide reglado».

(*) Según veremos más adelante, de la curva S se puede decir, no sólo que no es hipérbola, sino que, además, es elipse.

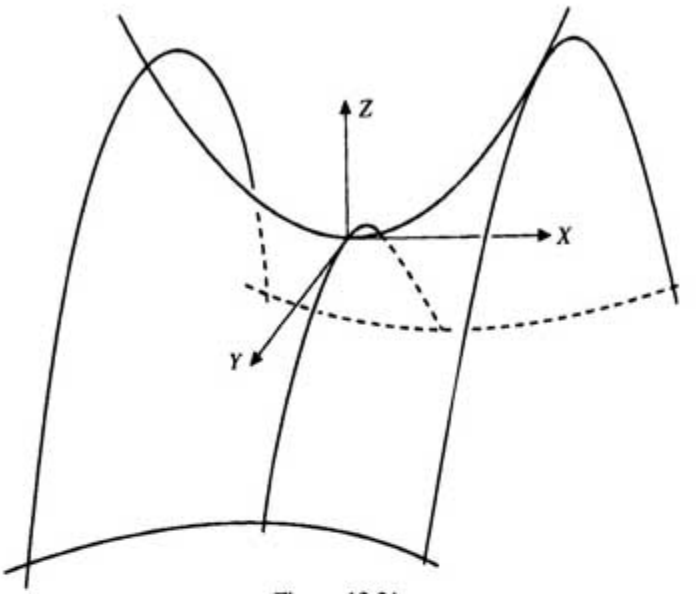


Figura 13.21.

COMPROBACION

Cuanto acabamos de decir, se comprueba de manera análoga (casi idéntica) a la de [293].

EJERCICIO

Considérese el paraboloide hiperbólico que, en ciertos ejes $Oxyz$ rectangulares, admite por ecuación reducida a

$$2z = \frac{x^2}{p} - \frac{y^2}{q}$$

Compruébese que ninguna de las secciones planas de dicho paraboloide es una elipse.

RESOLUCION

Procediendo exactamente igual que con el ejercicio anterior, la curva S_z es ahora

$$S_z: \begin{cases} \dfrac{x^2}{p} - \dfrac{y^2}{q} = 2ax + 2by + 2c \\ z = 0 \end{cases}$$

Como esta curva es una hipérbola (por ser $p > 0$ y $q > 0$), la curva S no puede ser elipse(*) pues al proyectarla tiene ramas infinitas.

(*) Según veremos más adelante, de la curva S se puede decir, no sólo que no es elipse, si no que, además, es hipébola.

13.4. PRIMERAS PROPIEDADES DE LAS CUADRICAS

No vamos a realizar aquí un estudio detallado de las cuádricas; sólo abordaremos sus propiedades más inmediatas y ello de manera superficial. Más adelante se analizarán estas cuestiones con mayor detenimiento.

INTERSECCION DE UNA CUADRICA Y UNA RECTA

[295]

La intersección de una cuádrica (elipsoide, hiperboloide o paraboloide) y una recta consta de dos puntos (propios o del infinito), de un solo punto (es decir, un punto doble o dos puntos «confundidos» en uno) o es vacía (dos «puntos imaginarios»). Se dice, respectivamente, que la recta es *secante*, *tangente* o *exterior* a la cuádrica.

NOTA: También pudiera ocurrir que la recta estuviera incluida en la cuádrica; en tal caso, la cuádrica se llama «reglada». De esta posibilidad nos ocuparemos más adelante; véase [298].

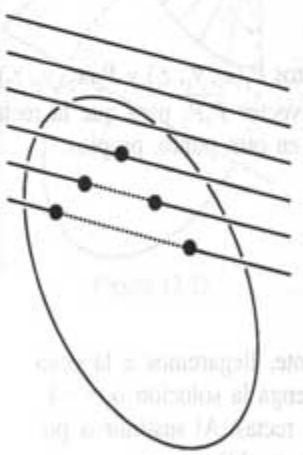


Figura 13.22.

COMPROBACION

Vamos a estudiar el caso del hiperboloide de una hoja; para las demás cuádricas se procede de modo análogo. Tomando unos ejes rectangulares *Oxyz* adecuados, el hiperboloide tendrá ecuación del tipo (reducida):

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \qquad [1]$$

Dada una recta cualquiera, si recurrimos a dos cualesquiera de sus puntos, (x_1, y_1, z_1) y (x_2, y_2, z_2), se la puede expresar mediante las siguientes ecuaciones paramétricas

$$x=\frac{\alpha x_1+x_2}{1+\alpha},\quad y=\frac{\alpha y_1+y_2}{1+\alpha},\quad z=\frac{\alpha z_1+z_2}{1+\alpha}\quad (\alpha\in\mathbb{R}:\ \alpha\neq -1)^{(*)} \qquad [2]$$

Llevando estas expresiones [2] a la ecuación [1] del elipsoide, nos encontramos

(*) Estas paramétricas se obtienen, a partir de las paramétricas usuales $x=x_1+\lambda(x_2-x_1)$, $y=y_1+\lambda(y_2-y_1)$, $z=z_1+\lambda(z_2-z_1)$, haciendo el cambio de parámetro $\lambda=1:(1+\alpha)$. Nótese que, aunque éstas paramétricas no proporcionan el punto (x_1, y_1, z_1), éste se obtiene como límite para $\alpha\to\infty$. Para $\alpha\to -1$, se obtiene (como posición límite) el punto del infinito de la recta.

con que los valores de α de los puntos de intersección son las raíces de la ecuación:

$$\left(\frac{x_1^2}{a^2}+\frac{y_1^2}{b^2}-\frac{z_1^2}{c^2}-1\right)\alpha^2+2\left(\frac{x_1x_2}{a^2}+\frac{y_1y_2}{b^2}-\frac{z_1z_2}{c^2}-1\right)\alpha+\left(\frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}-\frac{z_2^2}{c^2}-1\right)=0 \quad [3]$$

Esta ecuación [3], de segundo grado en la incógnita α, tiene dos soluciones reales distintas (recta secante), una solución real doble (recta tangente) o dos soluciones complejas (recta exterior)(*).

EJERCICIO

Considérese el hiperboloide de una hoja que, respecto de ciertos ejes rectangulares $Oxyz$, tiene ecuación (reducida):

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

Se considera también la recta que pasa por los puntos $P_1(x_1, y_1, z_1)$ y $P_2(x_2, y_2, z_2)$. Obtener la condición que tiene que verificar el vector $\overrightarrow{P_1P_2}$ para que la recta corte al hiperboloide en su punto del infinito (y en otro punto, propio).

RESOLUCION

Procediendo como en la comprobación precedente, llegaremos a la ecuación [3]. Ahora hay que imponer que esta ecuación tenga la solución $\alpha=-1$ (que es la que proporciona el punto del infinito de la recta). Al sustituir α por -1 en [3] y reagrupar adecuadamente los sumandos, se obtiene

$$\frac{(x_2-x_1)^2}{a^2}+\frac{(y_2-y_1)^2}{b^2}-\frac{(z_2-z_1)^2}{c^2}=0 \quad \text{o} \quad \frac{u^2}{a^2}+\frac{v^2}{b^2}-\frac{w^2}{c^2}=0$$

que es la relación pedida (u, v y w son las coordenadas del vector $\overrightarrow{P_1P_2}$). Nótese que (según lo dicho en el ejercicio de [290]), la relación obtenida significa que la recta dada ha de tener la dirección de una de las generatrices del cono asintótico del hiperboloide.

(*) Si los tres coeficientes de la ecuación [3] se anulasen, todos los puntos de la recta verificarían a la ecuación de la cuádrica y, por ello, la recta estaría incluida en la cuádrica. Esta posibilidad se analiza más adelante, en [298].

CONO TANGENTE DESDE UN PUNTO

[296]

En el espacio euclídeo, se considera una cuádrica cualquiera, esto es, o con centro (elipsoide o hiperboloide de una o dos hojas) o paraboloide (elíptico o hiperbólico), cuya ecuación reducida es, pues, del tipo:

$$\frac{x^2}{a^2}\pm\frac{y^2}{b^2}\pm\frac{z^2}{c^2}=1 \quad \text{(con centro)} \quad \text{o} \quad 2z=\frac{x^2}{p}\pm\frac{y^2}{q} \quad \text{(paraboloide)}$$

Si $P_0(x_0, y_0, z_0)$ es un punto que no pertenece a la cuádrica, las tangentes a ésta desde P_0 forman un cono cuadrático, que se llama *cono tangente*, cuya ecuación es:

$$\left(\frac{xx_0}{a^2}\pm\frac{yy_0}{b^2}\pm\frac{zz_0}{c^2}-1\right)^2=\left(\frac{x^2}{a^2}\pm\frac{y^2}{b^2}\pm\frac{z^2}{c^2}-1\right)\left(\frac{x_0^2}{a^2}\pm\frac{y_0^2}{b^2}\pm\frac{z_0^2}{c^2}-1\right)$$

(para las cuádricas con centro)

[1]

$$\left(z+z_0-\frac{xx_0}{p}\mp\frac{yy_0}{q}\right)^2=\left(2z-\frac{x^2}{p}\mp\frac{y^2}{q}\right)\left(2z_0-\frac{x_0^2}{p}\mp\frac{y_0^2}{q}\right)$$

(para los paraboloides)

Los correspondientes puntos de tangencia son los de intersección de la cuádrica con el siguiente plano, que se llama *plano polar* de P_0:

$$\frac{xx_0}{a^2}\pm\frac{yy_0}{b^2}\pm\frac{zz_0}{c^2}=1 \quad \text{(para las cuádricas con centro)}$$

[2]

$$z+z_0=\frac{xx_0}{p}\pm\frac{yy_0}{q} \quad \text{(para los paraboloides)}$$

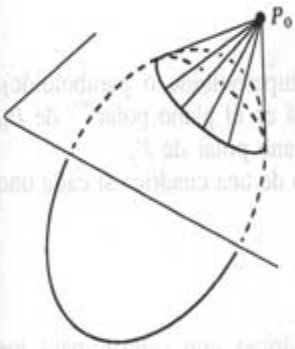


Figura 13.23.

DEMOSTRACION

Nos ocuparemos, por ejemplo, del caso del paraboloide elíptico; en los restantes casos se razonaría de modo similar.

Dada una recta cualquiera, la que pasa por dos puntos (x_1, y_1, z_1) y (x_2, y_2, z_2) distintos, expresémosla mediante las ecuaciones paramétricas [2] de [295] y hallemos su intersección con el paraboloide $2z = (x^2/p) + (y^2/q)$, como se hizo en [295] con el hiperboloide, lo que nos conduce a la ecuación (en la incognita α):

$$\left(2z_1-\frac{x_1^2}{p}-\frac{y_1^2}{q}\right)\alpha^2+2\left(z_1+z_2-\frac{x_1x_2}{p}-\frac{y_1y_2}{q}\right)\alpha+\left(2z_2-\frac{x_2^2}{p}-\frac{y_2^2}{q}\right)=0$$

cuyas raíces proporcionan los puntos de intersección. La recta será, pues, tangente al paraboloide si esta ecuación tiene una raíz doble, es decir, si su discriminante es nulo, o sea, si:

$$\left(z_1+z_2-\frac{x_1x_2}{p}-\frac{y_1y_2}{q}\right)^2=\left(2z_1-\frac{x_1^2}{p}-\frac{y_1^2}{q}\right)\left(2z_2-\frac{x_2^2}{p}-\frac{y_2^2}{q}\right) \qquad [A]$$

Las tangentes al paraboloide que pasan por el punto (x_0, y_0, z_0) están formadas por los puntos (x, y, z) tales que la recta que los une con (x_0, y_0, z_0) es tangente al paraboloide, es decir, tales que se verifica [A] para (x_0, y_0, z_0) y (x, y, z), en lugar de (x_1, y_1, z_1) y (x_2, y_2, z_2), y ésta relación es, justamente la [1] del enunciado, la que había que obtener. Para concluir totalmente la demostración habrá que comprobar que las rectas que forman este «cono tangente», que pasan todas por P_0, se obtienen de proyectar (desde P_0) los puntos de la cuádrica que están en una cónica situada en un plano que no pase por P_0. Como las cónicas sólo las conocemos, aún, a través de sus ecuaciones reducidas, dicha comprobación nos iba a resultar muy laboriosa, por lo que la aplazaremos para más adelante (*).

Los puntos de tangencia, de las tangentes desde P_0 a la cuádrica, son aquellos que están en el «cono tangente» y en la cuádrica, esto es, aquellos que verifican a la ecuación de la cuádrica y a la relación [1]. Como, por verificarse la ecuación de la cuádrica, en la relación [1] el primer factor del segundo miembro es nulo, resulta que los puntos de tangencia son los de la cuádrica que anulan al primer miembro de [1], esto es, los de la cuádrica que están en el plano que se ha llamado, en el enunciado, plano polar de P_0.

EJERCICIO

Considérese una cuádrica cualquiera (elipsoide, hiperboloide o paraboloide). Dados dos puntos P_1 y P_2, pruébese que si P_2 está en el plano polar(**) de P_1, respecto de la cuádrica, entonces, P_1 está en el plano polar de P_2.

Nota: Dos puntos se dicen *conjugados* respecto de una cuádrica si cada uno de ellos está en el plano polar del otro.

RESOLUCION

Consideremos, por ejemplo, el caso de las cuádricas con centro; para los paraboloides se procede de igual manera. Los planos polares de $P_1(x_1, y_1, z_1)$ y de $P_2(x_2, y_2, z_2)$ son, respectivamente (véase [296], [2]):

$$\frac{xx_1}{a^2} + \frac{yy_1}{b^2} + \frac{zz_1}{c^2} = 1 \quad \text{y} \quad \frac{xx_2}{a^2} + \frac{yy_2}{b^2} + \frac{zz_2}{c^2} = 1$$

El punto P_2 pertenecerá al plano polar de P_1 si se verifica que

$$\frac{x_2x_1}{a_2} + \frac{y_2y_1}{b^2} + \frac{z_2z_1}{c^2} = 1$$

Esta relación es, evidentemente, la que expresa, también, que el punto P_1 está en el plano polar de P_2, con lo que queda probada la propiedad del enunciado.

(*) Sobre todo esto se volverá, más adelante (véase [329]), cuando se haga el estudio general de las cuádricas.

(**) La ecuación del plano polar está dada en [296], [2]. Si el punto es de la cuádrica, el plano que tiene esta ecuación es también, como veremos enseguida, el plano tangente a la cuádrica en el susodicho punto.

PLANO TANGENTE EN UN PUNTO

[297]

En el espacio euclídeo, se considera una cuádrica cualquiera (con centro o paraboloide), cuya ecuación reducida es, pues, del tipo:

$$\frac{x^2}{a^2} \pm \frac{y^2}{b^2} \pm \frac{z^2}{c^2} = 1 \quad \text{(con centro)} \qquad 2z = \frac{x^2}{p} \pm \frac{y^2}{q} \quad \text{(paraboloide)}$$

Si $P_0(x_0, y_0, z_0)$ es un punto de la cuádrica, las tangentes a está en P_0 forman un plano (*plano tangente*), cuya ecuación es

$$\frac{xx_0}{a^2} \pm \frac{yy_0}{b^2} \pm \frac{zz_0}{c^2} = 1 \quad \text{(para las cuádricas con centro)} \qquad [1]$$

$$z + z_0 = \frac{xx_0}{p} \pm \frac{yy_0}{q} \quad \text{(para las paraboloides)}$$

La intersección de la cuádrica con su plano tangente en P_0 es (para cualquiera que sea el punto P_0 de ella):

- Un solo punto (dos rectas imaginarias que se cortan en un punto real) para los elipsoides, los hiperboloides de dos hojas (o elípticos) y los paraboloides elípticos; para expresar este hecho, se dice que estas tres cuádricas tienen todos sus puntos elípticos.
- Un par de rectas (que pasan por P_0) para los hiperboloides de una hoja (o hiperbólicos) y los paraboloides hiperbólicos; para expresar este hecho se dice que éstas cuádricas tienen todos sus puntos hiperbólicos.

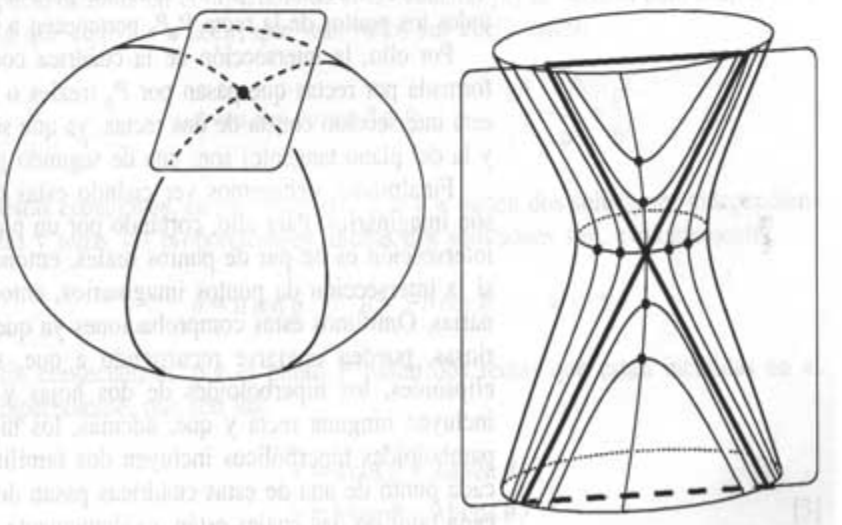


Figura 13.24.

DEMOSTRACION

Todo el razonamiento seguido en la demostración de [296] para comprobar que las tangentes desde P_0 a la cuádrica engendraban el lugar geométrico cuya ecuación era la [1] de [296], es válido también ahora. Pero como aquí suponemos que P_0 es de la cuádrica, resulta que el último factor de la referida ecuación [1] es, aquí, nulo, con lo que la [1] de [296] se reduce a la [1] nuestra, que es la ecuación de un plano. En él están, pues, incluidas todas las tangentes a la cuádrica en el punto P_0 de ella.

Veamos ahora que si la cuádrica y su plano tangente en P_0 tienen, además del P_0, otro punto común, $P_1(x_1, y_1, z_1)$, entonces todos los puntos de la recta P_0P_1 pertenecen, también, a la cuádrica y al plano tangente. En efecto: procediendo como en la comprobación de [295], un punto genérico $P(x, y, z)$ de la recta P_0P_1 se puede poner:

$$x = \frac{\alpha x_0 + x_1}{1+\alpha}, \quad y = \frac{\alpha y_0 + y_1}{1+\alpha}, \quad z = \frac{\alpha z_0 + z_1}{1+\alpha} \quad (\text{para } \alpha \in \mathbb{R}, \text{ con } \alpha \neq -1) \quad [2]$$

y la intersección de dicha recta con la cuádrica (que suponemos es de las con centro; para las paraboloides se procede de igual modo) se obtiene para los valores de α que hacen que las relaciones [2] verifiquen a la ecuación de la cuádrica, esto es, para los $\alpha \in \mathbb{R}$ que verifican a:

$$\left(\frac{x_0^2}{a^2} \pm \frac{y_0^2}{b^2} \pm \frac{z_0^2}{c^2} - 1\right)\alpha^2 + 2\left(\frac{x_0x_1}{a_2} \pm \frac{y_0y_1}{b^2} \pm \frac{z_0z_1}{c^2} - 1\right)\alpha + \left(\frac{x_1^2}{a^2} \pm \frac{y_1^2}{b^2} \pm \frac{z_1^2}{c^2} - 1\right) = 0 \quad [3]$$

Ahora bien, en esta ecuación todos los coeficientes son nulos: el de α^2 es nulo porque P_0 es de la cuádrica; el de α es cero porque P_1 está en el plano tangente en P_0; el término independiente es nulo porque P_1 está en la cuádrica. Por tanto, la ecuación [3] se satisface para todo $\alpha \in \mathbb{R}$, lo que significa que todos los puntos de la recta P_0P_1 pertenecen a la cuádrica y al plano tangente.

Por ello, la intersección de la cuádrica con su plano tangente en P_0 está formada por rectas que pasan por P_0 (reales o imaginarias); más exactamente, esta intersección consta de dos rectas, ya que sus ecuaciones (la de la cuádrica y la del plano tangente) son: una de segundo grado y la otra de primero.

Finalmente, deberemos ver cuándo estas dos rectas son reales y cuándo son imaginarias. Para ello, cortando por un plano que no contenga a P_0, si la intersección es un par de puntos reales, entonces las dos rectas son reales, y si la intersección da puntos imaginarios, entonces las dos rectas son imaginarias. Omitimos estas comprobaciones ya que, además de ser bastante laboriosas, pueden obviarse recurriendo a que, según veremos enseguida, los elipsoides, los hiperboloides de dos hojas y los paraboloides elípticos no incluyen ninguna recta y que, además, los hiperboloides de una hoja y los paraboloides hiperbólicos incluyen dos familias de rectas, de modo que por cada punto de una de estas cuádricas pasan dos de las citadas rectas (una de cada familia), las cuales están, evidentemente, en el plano tangente en dicho punto.

EJERCICIO

Considérese el hiperboloide de una hoja que, en ciertos ejes rectangulares $Oxyz$, admite por ecuación reducida a:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$

Tómese un punto P, cualquiera, de la elipse de garganta del hiperboloide (esto es, su elipse situada en $z=0$) y pruébese que por P pasan dos rectas que están incluidas en el hiperboloide; hállense dichas rectas y compruébese que pertenecen al plano tangente al hiperboloide en P.

RESOLUCION

El punto P tendrá coordenadas $P(a\cos\theta,\ b\,\text{sen}\,\theta,\ 0)$, para cierto $\theta \in [0,\ 2\pi[$. Una recta que pase por P se podrá expresar, en función de un vector $(u,\ v,\ w)$ de su dirección, mediante sus ecuaciones paramétricas:

$$(x,\ y,\ z)=(a\cos\theta,\ b\,\text{sen}\,\theta,\ 0)+\lambda(u,\ v,\ w) \qquad [1]$$

Al llevar estas expresiones de x, y y z a la ecuación del hiperboloide se obtiene la siguiente ecuación, en la incógnita λ:

$$\lambda^2\left(\frac{u^2}{a^2}+\frac{v^2}{b^2}-\frac{w^2}{c^2}\right)+2\lambda\,\frac{1}{ab}\,(ub\cos\theta+va\,\text{sen}\,\theta)+0=0 \qquad [2]$$

Las raíces de esta ecuación llevadas a [1] dan las coordenadas de los puntos de intersección de la recta y el hiperboloide (una de esas raíces es $\lambda=0$, ya que para $\lambda=0$ se obtiene el punto P, que es del hiperboloide). La recta estará, pues, incluida en el hiperboloide si la ecuación [2] se verifica para todo $\lambda \in \mathbb{R}$, lo que equivale a decir, que sean nulos sus coeficientes:

$$ub\cos\theta+va\,\text{sen}\,\theta=0 \quad \text{y} \quad \frac{w^2}{c^2}=\frac{u^2}{a^2}+\frac{v^2}{b^2}$$

Estas ecuaciones, en las incógnitas u, v y w tienen dos soluciones independientes y todas sus proporcionales; dichas dos soluciones son, evidentemente:

$$u=a\,\text{sen}\,\theta \quad , \quad v=-b\cos\theta \quad , \quad w=\pm c$$

En consecuencia, por el punto P pasan dos rectas que están incluidas en el hiperboloide, que son las:

$$\left.\begin{aligned} x &= a\cos\theta \mp \lambda a\,\text{sen}\,\theta \\ y &= b\,\text{sen}\,\theta \mp \lambda b\cos\theta \\ z &= \qquad \pm \lambda c \end{aligned}\right\} \qquad [3]$$

Finalmente, de acuerdo con [297], la ecuación del plano tangente al hiperboloide en $P\ (a\cos\theta,\ b\,\text{sen}\,\theta,\ b)$ es:

$$\frac{x\cos\theta}{a}+\frac{y\,\text{sen}\,\theta}{b}=1 \qquad [4]$$

Llevando las ecuaciones [3] a la [4], se obtiene una identidad, para todo λ, lo que confirma que las rectas [3] están incluidas en el plano tangente [4].

CUADRICAS REGLADAS

[298]

El hiperboloide de una hoja (o hiperbólico) y el paraboloide hiperbólico son *cuádricas regladas*; es decir: para cada una de estas cuádricas hay dos familias uniparamétricas de rectas que están, todas, incluidas en la cuádrica. Estas familias se llaman «sistemas de generatrices rectilíneas». Por cada punto de una de estas cuádricas pasan dos generatrices rectilíneas: una de cada sistema; en consecuencia, todos los puntos de estas cuádricas son hiperbólicos. Las generatrices rectilíneas de las cuádricas regladas son:

Cuádrica	Generatrices rectilíneas
$\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1$ (hiperboloide hiperbólico)	$\dfrac{x-a\cos\theta}{a\,\text{sen}\,\theta}=\dfrac{y-b\,\text{sen}\,\theta}{-b\cos\theta}=\dfrac{z}{\pm c}$ o $\begin{cases} x=a\left(\cos\theta\pm\dfrac{\text{sen}\,\theta}{c}z\right) \\ y=b\left(\text{sen}\,\theta\mp\dfrac{\cos\theta}{c}z\right) \end{cases}$ (θ parámetro; $0\leq\theta<2\pi$)
$2z=\dfrac{x^2}{p}-\dfrac{y^2}{q}$ (paraboloide hiperbólico)	$\dfrac{x-\sqrt{p}\,\mu}{\sqrt{p}}=\dfrac{\pm y+\sqrt{q}\,\mu}{\sqrt{q}}=\dfrac{z}{2\mu}$ o $\begin{cases} x=\sqrt{p}\left(\mu+\dfrac{z}{2\mu}\right) \\ y=\pm\sqrt{q}\left(\mu-\dfrac{z}{2\mu}\right) \end{cases}$ ($\mu\in\mathbb{R}$ parámetro)(*)

(*) Para $\mu\to 0$, se obtienen las generatrices rectilíneas situadas en $z=0$, que son las $x\sqrt{p}=\pm y\sqrt{q},\ z=0$.

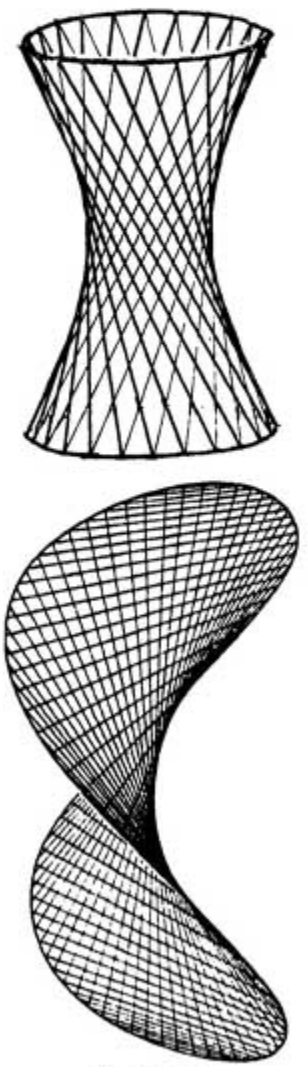


Figura 13.25.

DEMOSTRACION

Consideremos el caso del hiperboloide elíptico; para el paraboloide elíptico se razona de igual modo. Buscamos rectas que estén incluidas en la cuádrica; nótese que ninguna de ellas puede ser paralela al plano $z = 0$ (pues, al hacer z constante en la ecuación del hiperboloide, se obtienen elipses). Así, pues, las rectas buscadas serán del tipo:

$$(x, y, z) = (\alpha, \beta, 0) + \lambda(u, v, w) \quad \lambda \in \mathbb{R} \text{ parámetro}$$

Como $(\alpha, \beta, 0)$ debe ser de la cuádrica, será un punto de su elipse de garganta, esto es, del tipo $\alpha = a\cos\theta$, $\beta = b\,\text{sen}\,\theta$ para algún $\theta \in [0, 2\pi[$. Prosiguiendo, a partir de aquí, como en el último ejercicio, se llega al resultado

$$x = a\cos\theta + \lambda a\,\text{sen}\,\theta \quad , \quad y = a\,\text{sen}\,\theta - \lambda b\cos\theta \quad , \quad z = \pm\lambda c$$

que coincide con lo dicho en el enunciado.

Para comprobar que por cada punto (x_0, y_0, z_0) de la cuádrica pasa una generatriz de cada una de las dos familias, consideremos el caso del paraboloide hiperbólico (un razonamiento análogo se puede hacer para el hiperboloide hiperbólico) y, de las dos familias de generatrices tomemos, por ejemplo, la que se obtiene para el signo más (en su ecuación, donde pone $\pm$, tómese $+$). Hay, pues, que probar que el sistema de ecuaciones (μ es la incógnita; x_0, y_0, z_0 son las coordenadas de un punto cualquiera de la cuádrica):

$$x_0 = \sqrt{p}\left(\mu + \frac{z_0}{2\mu}\right) \quad , \quad y_0 = \sqrt{q}\left(\mu - \frac{z_0}{2\mu}\right) \text{ (*)}$$

es compatible y tiene solución única. Este sistema es, evidentemente, equivalente al:

$$\frac{x_0}{\sqrt{p}} + \frac{y_0}{\sqrt{q}} = 2\mu \quad , \quad \frac{x_0}{\sqrt{p}} - \frac{y_0}{\sqrt{q}} = \frac{z_0}{\mu}$$

Ahora bien, estas dos ecuaciones son compatibles (tienen la misma solución para μ), si y sólo si se verifica que:

$$\left(\frac{x_0}{\sqrt{p}} + \frac{y_0}{\sqrt{q}}\right)\left(\frac{x_0}{\sqrt{p}} - \frac{y_0}{\sqrt{q}}\right) = 2z_0, \quad \text{sea} \quad \frac{x_0^2}{p} - \frac{y_0^2}{q} = 2z_0$$

Como esta relación es cierta, pues (x_0, y_0, z_0) es de la cuádrica, el sistema es, en efecto, compatible y conduce a una, única, generatriz (de la familia elegida) que pasa por el punto dado.

OBSERVACION

Sabiendo ya que existen, hay otro modo, quizá más fácil, de obtener las generatrices rectilíneas de las cuádricas regladas, que conduce al siguiente resultado:

(*) Suponemos $\mu \neq 0$. Para el caso $\mu = 0$ la propiedad también se verifica, como se comprueba trivialmente.

[299]

Las generatrices rectilíneas de las cuádricas regladas son:

Cuádrica	Generatrices rectilíneas
$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$ (hiperboloide hiperbólico)	$\begin{cases}\frac{x}{a}+\frac{z}{c}=\lambda\left(1\pm\frac{y}{b}\right)\\ \lambda\left(\frac{x}{a}-\frac{z}{c}\right)=\lambda\left(1\mp\frac{y}{b}\right)\end{cases}$
$2z=\frac{x^2}{p}-\frac{y^2}{q}$ (paraboloide hiperbólico)	$\begin{cases}\frac{x}{\sqrt{p}}\pm\frac{y}{\sqrt{q}}=2\lambda\\ \lambda\left(\frac{x}{\sqrt{p}}\mp\frac{y}{\sqrt{q}}\right)=z\end{cases}$

Para cada valor de λ se obtiene una generatriz de cada una de las familias. Para $\lambda\to\infty$ se obtiene, como posición límite, una generatriz cuya ecuación es la que resulta de poner $\lambda=1/\mu$, quitar denominadores y, finalmente, hacer $\mu=0$.

COMPROBACION

Expresemos las ecuaciones de las dos cuádricas del siguiente modo:

Hiperboloide hiperbólico: $\left(\frac{x}{a}+\frac{z}{c}\right)\left(\frac{x}{a}-\frac{z}{c}\right)=\left(1+\frac{y}{b}\right)\left(1-\frac{y}{b}\right)$

Paraboloide hiperbólico: $\left(\frac{x}{\sqrt{p}}+\frac{y}{\sqrt{p}}\right)\left(\frac{x}{\sqrt{p}}-\frac{y}{\sqrt{p}}\right)=2z$

Todo punto de una cualquiera de las rectas del enunciado, por verificar a sus dos ecuaciones, verifica a la ecuación que resulta de multiplicar, miembro a miembro, dichas dos ecuaciones y, como lo que así se obtiene es la ecuación de la cuádrica, resulta que todos los puntos de dichas rectas pertenecen a la cuádrica, como había que comprobar.

EJERCICIO

Considérese el paraboloide hiperbólico que, en ejes rectangulares $Oxyz$, admite por ecuación reducida a

$$2z=\frac{x^2}{p}-\frac{y^2}{q}\quad(p>0\text{ y }q>0)$$

Hallar el lugar geométrico que forman los puntos de esta cuádrica en los que las dos generatrices rectilíneas son perpendiculares.

RESOLUCION

Sea $P(x, y, z)$ un punto genérico del lugar pedido. Este punto estará en las generatrices (una de cada familia) que se obtengan para ciertos parámetros μ y μ' (véase [298]), esto es, debe ser:

$$\begin{cases} x = \sqrt{p}\left(\mu + \dfrac{z}{2\mu}\right) \\ y = \sqrt{q}\left(\mu - \dfrac{z}{2\mu}\right) \end{cases} \quad \text{y} \quad \begin{cases} x = \sqrt{p}\left(\mu' + \dfrac{z}{2\mu'}\right) \\ y = -\sqrt{q}\left(\mu' - \dfrac{z}{2\mu'}\right) \end{cases} \qquad [1]$$

Como estas rectas han de ser perpendiculares y tienen las direcciones de los vectores $(\sqrt{p}, \sqrt{q}, 2\mu)$ y $(\sqrt{p}, \sqrt{q}, 2\mu')$, se ha de verificar que

$$p + q + 4\mu\mu' = 0 \qquad [2]$$

Eliminando los parámetros μ y μ' entre las ecuaciones [1] y [2], se obtienen dos relaciones(*) entre x, y y z, que son las del lugar pedido. Las ecuaciones [1] son equivalentes a las

$$\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} = 2\mu = \frac{z}{\mu'} \quad \text{y} \quad \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = \frac{z}{\mu} = 2\mu' \qquad [3]$$

Eliminando, pues, μ y μ' entre [2] y [3], se obtiene:

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z \quad \text{y} \quad z = -\frac{1}{2}(p + q)$$

Por tanto, el lugar pedido lo forman los puntos del paraboloide que están en el plano $z = -(p + q)/2$, que es una hipérbola:

$$\frac{x^2}{p} - \frac{y^2}{q} = -(p + q), \quad z = -\frac{1}{2}(p + q)$$

CUADRICAS NO REGLADAS

[300]

El elipsoide, el hiperboloide de dos hojas (o elíptico) y el paraboloide elíptico son cuádricas que no contienen rectas (ni una sola), por lo que se llaman cuádricas no regladas. Todos los puntos de estas cuádricas son elípticos.

(*) Como hay 5 ecuaciones, al eliminar 2 parámetros, parece que debieran quedar 3 ecuaciones (y no 2 como se ha dicho). No obstante, como las dos generatrices rectilíneas se cortan, en (x, y, z), las relaciones [1] no son independientes, sino que existe una relación entre ellas, por lo que, a la postre se van a obtener sólo 2 ecuaciones independientes en x, y y z.

COMPROBACION

Si una cuádrica incluye a una recta, el punto de corte de esta recta con un plano sería de la cuádrica. Si al cortar una cuádrica por un plano se obtuviera intersección vacía, la cuádrica no podría contener ninguna recta, salvo quizá las rectas paralelas al plano. Si al cortar una cuádrica por tres planos, π_1, π_2 y π_3, que forman triedro, se obtiene intersección vacía, entonces la cuádrica no incluye ninguna recta, no es reglada. Para cada una de las cuádricas del enunciado, es evidente que los siguientes planos π_1, π_2 y π_3 forman triedro y dan intersección vacía con la cuádrica, por lo que las referidas cuádricas son, las tres, no regladas:

1. Para el elipsoide E:

$$E: \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} \quad ; \quad \pi_1: x=2a \quad , \quad \pi_2: y=2b \quad , \quad \pi_3: z=2c$$

2. Para el hiperboloide elíptico H:

$$H: \frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad ; \quad \pi_1: x=0 \quad , \quad \pi_2: \frac{x}{a}=\frac{y}{b} \quad , \quad \pi_3: \frac{x}{a}=\frac{z}{c}$$

3. Para el paraboloide elíptico P:

$$P: 2z=\frac{x^2}{p}+\frac{y^2}{q} \ (p>0,\ q>0) \quad ;$$

$$\pi_1: z=-1 \ , \ \pi_2: x=pz+1 \ . \ \pi_2: y=2q+1$$

CAPITULO

14

Cónicas y cuádricas: Análisis afín y métrico

Hasta este momento, en el capítulo anterior, nos hemos ocupado en analizar «las tres cónicas» y «las cinco cuádricas» y ello se ha hecho a partir de sus ecuaciones reducidas, las cuales han sido, en todos los casos, de segundo grado en las coordenadas x e y (en el caso de las cónicas) o en las x, y y z (para las cuádricas). Es llegado el momento de estudiar todas aquellas curvas del plano y superficies del espacio que tienen ecuación de segundo grado, a las que se llama variedades (es decir, curvas o superficies) cuadráticas o de segundo orden, viendo en primer lugar si, además de las cónicas y las cuádricas, hay más tipos de dichas variedades. Nos vamos a encontrar con que las elipses, parábolas e hipérbolas son las únicas cónicas «ordinarias» o «regulares», pero que también hay curvas cuadráticas «degeneradas» o «singulares», que son los pares de rectas (distintas o coincidentes); en el caso de las cuádricas, además de los elipsoides, paraboloides e hiperboloides, que son las cuádricas «ordinarias», hay otras superficies cuadráticas, llamadas cuádricas «degeneradas», que son conos, cilindros y pares de planos (distintos o confundidos).

Dada una variedad (curva o superficie) cuadrática, para llegar a obtener su clasificación y su ecuación reducida, vamos a tener que estudiar cómo varía su ecuación con los cambios de referencia cartesiana; este análisis nos llevará a obtener los «invariantes métricos»: expresiones, obtenidas a partir de los coeficientes de la ecuación de la variedad, que permanecen fijos o invariantes frente a los cambios de referencia rectangular.

α LAS CONICAS

En el plano afín, las curvas de segundo orden son los lugares geométricos que, respecto de ejes cartesianos Oxy, admiten ecuación del tipo:

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0$$

para ciertos números reales a, b, c, d, e y f (coeficientes), donde a, b y c no podrán ser los tres nulos. Si a las coordenadas se las denota por x_1 y x_2 y se recurre también a los subíndices para designar a los coeficientes, la anterior ecuación se podrá expresar de cualquiera de las formas siguientes, que en realidad son la misma:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0$$

$$\sum_{i,j=1}^{2} a_{ij}x_ix_j + 2\sum_{h=1}^{2} b_hx_h + c = 0 \quad (a_{21} = a_{12})$$

$$[x_1\ x_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 2[b_1\ b_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [c] = O$$

$$[1x_1\ x_2]\left[\begin{array}{c|cc} c & b_1 & a_{11} \\ \hline b_1 & a_{12} & b_2 \\ b_2 & a_{21} & a_{22} \end{array}\right]\begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} = O \quad \text{o} \quad [1\ X^t]\begin{bmatrix} c & B^t \\ B & A \end{bmatrix}\begin{bmatrix} 1 \\ X \end{bmatrix} = O$$

14.1. LAS CONICAS: ECUACION Y PRIMERAS PROPIEDADES

DEFINICION DE CONICA

[301]

En el plano afín, se llama cónica al lugar geométrico de los puntos cuyas coordenadas (x_1, x_2), respecto de una referencia cartesiana, verifican a una ecuación del tipo

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0$$

$$\text{o} \quad \sum_{i,j=1}^{2} a_{ij}x_ix_j + 2\sum_{h=1}^{2} b_hx_h + c = 0 \qquad [1]$$

para ciertos a_{ij}, b_h, $c \in \mathbb{R}$ y donde a_{11}, $a_{12} = a_{21}$ y a_{22} no son los tres nulos. Con notación matricial, recurriendo a las matrices $A = [a_{ij}]$ y $B = [b_h]$ y llamando X a la columna de las coordenadas x_1 y x_2, la ecuación [1] se puede poner:

$$X^tAX + 2B^tX + [c] = O \quad \text{o} \quad [1\ X^t]M\begin{bmatrix} 1 \\ X \end{bmatrix} = O \quad \text{con} \quad M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \qquad [2]$$

A la matriz M, o a cualquiera de sus proporcionales, se le llama *matriz de la cónica* en los ejes dados; a la matriz A la llamaremos *matriz de los términos cuadráticos*.

Nota: En breve (véase [304]) comprobaremos que la ecuación de una cónica es de segundo grado en cualquier referencia cartesiana del plano.

EJEMPLOS

La anterior ecuación general de una cónica da acogida a «curvas» de muy distinto tipo. En primer lugar habrá que citar a las elipses, parábolas e hipérbolas, que conocemos a través de sus ecuaciones reducidas (a las coordenadas las llamamos ahora x e y):

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad y^2 = 2px$$

Pero no son éstas las únicas cónicas; también están las parejas de rectas y algunos otros casos más «patológicos» aún, como los siguientes:

$x^2 - y^2 = 0$ (pareja de rectas: $x + y = 0$ y $x - y = 0$)
$x^2 - 4 = 0$ (pareja de rectas: $x + 2 = 0$ y $x - 2 = 0$)
$x^2 = 0$ (la recta $x = 0$ doble, «contada dos veces»)
$x^2 + 4 = 0$ (conjunto vacío o «par de rectas imaginarias» $x \pm 2i = 0$)
$x^2 + y^2 = 0$ (el punto $x = 0$, $y = 0$ o las «rectas imaginarias» $y \pm ix = 0$)

MATRICES DE UNA CONICA

Para evitar situaciones indeseables, como el llegar a decir que las cónicas $x^2 + y^2 + 1 = 0$ y $x^2 + 4 = 0$ son la misma, por tratarse en ambos casos del conjunto vacío, se convendrá en que una cónica queda caracterizada (en los ejes que se estén considerando) por la matriz M de su ecuación o por cualquier otra matriz proporcional a M, ya que con ella se obtiene una ecuación equivalentes. Más exactamente:

[302]

En el plano afín y respecto de una referencia cartesiana, una cónica queda caracterizada por una matriz simétrica M, de tamaño 3×3, que tiene no nula la submatriz que resulta de prescindir de sus primeras fila y columna (llamada matriz de los términos cuadráticos). La cónica es equivalente al conjunto que forman todas las matrices (no nulas) proporcionales a M, de modo que:

$$\left(\text{cónica de ecuación}^{(*)}: \quad [1\ X^t]M\begin{bmatrix}1\\X\end{bmatrix} = O\right) \simeq \{\lambda M / \lambda \in \mathbb{R},\ \lambda \neq 0\}$$

Dos cónicas son iguales si, y sólo si, sus matrices (en la misma referencia cartesiana) son proporcionales.

(*) Por X se denota a la matriz columna de las coordenadas x_1, x_2 de un punto del plano.

EJERCICIO

Compruébese que si una cónica tiene dos puntos (distintos), entonces tiene infinitos puntos. (Esta propiedad, cuando hayamos avanzado un poco en el estudio de las cónicas, será algo obvio.)

RESOLUCION

Tomemos unos nuevos ejes que tengan su origen en uno de dichos puntos y cuyo «eje de la x» pase por el otro. En ellos, la ecuación de la cónica será del tipo

$$ax^2 + 2bxy + cy + 2dx + 2ey + f = 0 \quad \text{con} \quad f = 0,\ a \neq 0 \text{ y } d \neq 0$$

($f = 0$ por pasar por el origen; $a \neq 0$ y $d \neq 0$ para que, al cortar por $y = 0$, la ecuación $ax^2 + 2dx = 0$ tenga una raíz real no nula). Cortando la cónica por la recta $y = mx$, se obtiene:

$$\begin{cases} y = mx \\ (a + 2bm + cm^2)x^2 + 2(d + em)x = 0 \end{cases}$$

Como para m suficientemente pequeño es

$$a + 2bm + cm^2 \neq 0 \text{ (pues } a \neq 0) \quad \text{y} \quad d + em \neq 0 \text{ (pues } d \neq 0)$$

resulta que la última ecuación tiene solución no nula para rectas $y = mx$ próximas al eje $y = 0$ (para m suficientemente pequeña); estas soluciones proporcionan, pues, nuevos puntos de la cónica.

CAMBIO DE COORDENADAS

[303]

En el plano afín, se considera la cónica que, en una cierta referencia cartesiana (coordenadas x_1, x_2), tiene a M por matriz, es decir, que admite por ecuación a

$$[1 \; X^t] M \begin{bmatrix} 1 \\ X \end{bmatrix} = O, \quad \text{donde} \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \qquad [1]$$

($A \neq O$ simétrica 2×2; B fila 1×2; c término independiente.)

Supongamos, también, que se adopta una nueva referencia cartesiana (cordenadas x_1', x_2') y que el correspondiente cambio de coordenadas es (véase [255]):

$$\begin{bmatrix} 1 \\ X \end{bmatrix} = Q^* \begin{bmatrix} 1 \\ X' \end{bmatrix} \quad \text{con} \quad Q^* = \begin{bmatrix} 1 & O \\ C & Q \end{bmatrix} \qquad [2]$$

(donde: C es la columna de coordenadas del nuevo origen; Q es la matriz 2×2 correspondiente al cambio de base.)

Entonces, en la nueva referencia, la cónica tiene por matriz a la M' siguiente, que es congruente con M (véase [114]):

$$M' = Q^{*t} M Q^* \quad \text{o} \quad M' \equiv \begin{bmatrix} C & B'^t \\ B' & A' \end{bmatrix} = \begin{bmatrix} 1 & C^t \\ O & Q^t \end{bmatrix} \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \begin{bmatrix} 1 & O \\ C & Q \end{bmatrix} \qquad [3]$$

Y, en particular, las matrices de los términos cuadráticos, A y A', también son congruentes, pues están relacionadas mediante:

$$A' = Q^t A Q$$

COMPROBACION

Como la correspondencia $[\overline{2}]$ entre unas y otras coordenadas es biyectiva, llevando esta expresión [2] a la ecuación [1] de la cónica, se obtiene otra ecuación de ella, en las nuevas coordenadas, que es:

$$\left(Q^*\begin{bmatrix}1\\X'\end{bmatrix}\right)^t M\left(Q^*\begin{bmatrix}1\\X'\end{bmatrix}\right)=O \quad \text{o} \quad [1\ X'^t](Q^{*t}MQ^*)\begin{bmatrix}1\\X'\end{bmatrix}=O$$

lo que significa que la matriz M' de la cónica en la nueva referencia es $M'=Q^{*t}MQ^*$. Finalmente, multiplicando por bloques en este segundo miembro, se obtiene que el bloque de lugar (2, 2) de este producto es Q^tAQ, luego $A'=Q^tAQ$.

INVARIANTES PROYECTIVO Y AFIN

[304]

En el plano afín, el hecho de que una curva sea cónica, esto es, con ecuación de segundo grado, es independiente de la referencia cartesiana en la que aquella se expresa. Aún más:

Supóngase que, en una cierta referencia cartesiana del plano, M es la matriz simétrica de una cónica y A es su submatriz (simétrica) de los términos cuadráticos. Supóngase que, en otra referencia, las referidas matrices son M' y A'. Entonces, se verifica que las signaturas de M y A son invariantes, es decir:

$$\text{Sig } M' = \text{Sig } (\pm M) \quad \text{en particular} \quad [\det M' = 0] \Leftrightarrow [\det M = 0]$$
$$\text{Sig } A' = \text{Sig } (\pm M) \quad \text{en particular} \quad [\det A' = 0] \Leftrightarrow [\det A = 0]$$

DEMOSTRACION

Si llamamos Q^* a la matriz del cambio de coordenadas y Q a la matriz correspondiente al cambio de base (como se ha hecho en [303]), resulta que $Q^{*t}MQ^*$ y Q^tAQ son proporcionales a M' y A', o sea, existe $\rho \neq 0$ tal que

$$M' = \rho Q^{*t}MQ^* \quad \text{y} \quad A' = \rho Q^tAQ \qquad [1]$$

Como Q^* y Q son matrices regulares. nos encontramos con que M' es congruente con ρM y con que A' es congruente con ρA (para un cierto $\rho \neq 0$). Llamando $h=\sqrt{|\rho|}$, con lo que $\rho = \pm h^2$, podemos poner:

$$M' = (hQ^*)^t(\pm M)(hQ^*) \quad \text{y} \quad A' = (hQ)^t(\pm A)(hQ)$$

luego M' y A' son respectivamente congruentes con M y A o con $-M$ y $-A$. Dado que las matrices congruentes tienen la misma signatura, resulta que las signaturas de M y $\pm M$ son iguales y también lo son las de A y $\pm A$.

Nótese que, si sig $M = (p, q)$, la condición det $M = 0$ equivale a $p + q < 3$ y, como sig M' vale (p, q) o (q, p), resulta que la relación $p + q < 3$ equivale a su vez a det $M' = 0$. Análogamente se comprueba la equivalencia de det $A = 0$ y det $A' = 0$.

Finalmente, al cambiar la referencia en el plano, la ecuación de una cónica, de segundo grado (o sea, con matriz A de términos cuadráticos no nula), sigue siendo, pues, de segundo grado, dado que sig $(\pm A)$ no varía.

EJERCICIO

En el plano afín y respecto de cierta referencia cartesiana $(O; \bar{e}_1, \bar{e}_2)$ (coordenadas x e y), se considera una cónica cuya ecuación es del tipo:

$$x^2 - y^2 + \alpha x + \beta y + \gamma = 0$$

Hallar unos vectores $\bar{u} = (u_1, u_2)$ y $\bar{v} = (v_1, v_2)$ tales que en la referencia $(O; \bar{u}, \bar{v})$ (coordenadas x' e y'), la ecuación de la cónica sea del tipo

$$x'y' + \alpha' x' + \beta' y' + \gamma' = 0$$

RESOLUCION

De acuerdo con [313], las matrices de los términos cuadráticos han de estar relacionadas (para un cierto $\rho \neq 0$) mediante

$$\begin{bmatrix} 0 & \rho \\ \rho & 0 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} = \begin{bmatrix} u_1^2 - u_2^2 & u_1 v_1 - u_2 v_2 \\ u_1 v_1 - u_2 v_2 & v_1^2 - v_2^2 \end{bmatrix}$$

luego los $\bar{u}$ y $\bar{v}$ han de verificar que:

$$u_1^2 = u_2^2 \quad , \quad v_1^2 = v_2^2 \quad , \quad u_1 v_1 - u_2 v_2 \neq 0$$

luego $\bar{u}$ y $\bar{v}$ han de ser proporcionales a $(1, 1)$ y $(1, -1)$.

ECUACION HOMOGENEA DE UNA CONICA

Para estudiar algunas de las cuestiones relacionadas con las cónicas, van a sernos de gran utilidad las coordenadas homogéneas (véase [237]). Este es el caso, por ejemplo, de la intersección de una cónica con una recta o de las clasificaciones de las cónicas en ordinarias y degeneradas y en función de sus puntos del infinito.

Recuérdese que, en el plano afín y respecto de una cierta referencia cartesiana, si las coordenadas cartesianas (absolutas) de un punto son (x_1, x_2), entonces sus sistemas de coordenadas homogéneas son todas las ternas:

$$(X_0, X_1, X_2) = (\rho, \rho x_1, \rho x_2) \quad \text{para } \rho \neq 0 \text{ cualquiera}$$

Llamando X y $\mathcal{X}$ a las matrices columna de las coordenadas absolutas y homogéneas, respectivamente, del punto, será:

$$\mathcal{X} = \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{X}_1 \\ \mathcal{X}_2 \end{bmatrix} = \rho \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} = \rho \begin{bmatrix} 1 \\ X \end{bmatrix} \quad \text{con } \rho \neq 0$$

[305]

En el plano afín y respecto de una referencia cartesiana (coordenadas absolutas x_1, x_2; coordenadas homogéneas $\mathcal{X}_0$, $\mathcal{X}_1$, $\mathcal{X}_2$), la cónica cuya ecuación (absoluta) es [1] tiene a [2] por ecuación homogénea, siendo:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0 \qquad [1]$$

$$a_{11}\mathcal{X}_1^2 + 2a_{12}\mathcal{X}_1\mathcal{X}_2 + a_{22}\mathcal{X}_2^2 + 2b_1\mathcal{X}_0\mathcal{X}_1 + 2b_2\mathcal{X}_0\mathcal{X}_2 + c\mathcal{X}_0^2 = 0 \qquad [2]$$

Esta ecuación [2], al expresarla matricialmente, queda:

$$[\mathcal{X}_0\mathcal{X}_1\mathcal{X}_2] \begin{bmatrix} c & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \mathcal{X}_0 \\ \mathcal{X}_1 \\ \mathcal{X}_2 \end{bmatrix} = O$$

o, de modo abreviado:

$$\mathcal{X}^t M \mathcal{X} = O$$

Nota: Respecto de la notación, es usual llamar $c = a_{00}$, $b_1 = a_{01} = a_{10}$ y $b_2 = a_{02} = a_{20}$, con lo que $M = [a_{ij}]$ (para $i, j = 0, 1, 2$) y la ecuación [2] queda $\Sigma a_{ij}\mathcal{X}_i\mathcal{X}_j = 0$ (la suma para $i, j = 0, 1, 2$).

OBSERVACION

La ecuación [2], que se obtiene de la [1] sin más que poner $x_1 = \mathcal{X}_1/\mathcal{X}_0$ y $x_2 = \mathcal{X}_2/\mathcal{X}_0$ y quitar denominadores, es una ecuación homogénea de segundo grado, en las tres variables $\mathcal{X}_0$, $\mathcal{X}_1$, $\mathcal{X}_2$. Sin embargo, una ecuación homogénea de segundo grado, en dichas variables, puede no representar a una cónica a causa de la condición, ya señalada en [301], que exige la no anulación de algunos de los coeficientes a_{11}, $a_{12} = a_{21}$ y a_{22}. Si estos tres coeficientes fuesen nulos, la ecuación [1] lo sería de una recta y la ecuación [2], que se podría poner:

$$\mathcal{X}_0(2b_1\mathcal{X}_1 + 2b_2\mathcal{X}_2 + c) = 0$$

es de segundo grado porque, además de la citada recta, también incluye a la recta del infinito ($\mathcal{X}_0 = 0$). Así, pues, al pedir que los coeficientes a_{11}, $a_{12} = a_{21}$ y a_{22} no sean todos nulos, se está excluyendo de entre las cónicas a las que podrían estar formadas por la unión de una recta (propia) y la recta del infinito.

INTERSECCION DE CONICA Y RECTA

[306]

En el plano afín ampliado (con los puntos del infinito) una cónica y una recta se cortan en dos, uno o ningún punto y se dice que la recta es, respectivamente *secante*, *tangente* o *exterior* a la cónica(*).

Supongamos que, en una cierta referencia cartesiana (coordenadas homogéneas X_0, X_1, X_2), la cónica y la recta admiten por ecuaciones homogéneas a las:

$$X^t M X = O, \quad \text{donde} \begin{cases} X = [X_i] \text{ es la columna de coordenadas } (3 \times 1) \\ M \text{ es la matriz (simétrica) de la cónica} \end{cases}$$

$$X = \lambda\alpha + \mu\beta, \quad \text{donde} \begin{cases} X, \ \alpha \text{ y } \beta \text{ son columnas de} \\ \text{coordenadas del punto genérico} \\ \text{y de dos puntos fijos de la recta} \end{cases}$$

Entonces los puntos de su intersección se obtienen para los (λ, μ), parámetros, que son soluciones de la ecuación:

$$(\alpha^t M \alpha)\lambda^2 + 2(\alpha^t M \beta)\lambda\mu + (\beta^t M \beta)\mu^2 = O \qquad [1]$$

La recta es tangente a la cónica si se verifica la relación:

$$(\alpha^t M \beta)^2 = (\alpha^t M \alpha)(\beta^t M \beta) \qquad [2]$$

(*) Cabe también la posibilidad («patológica») de que la recta esté incluida en la cónica, cosa que, según veremos pronto, puede acontecer sólo si la cónica es degenerada. La recta está incluida en la cónica sí, y sólo si, los tres coeficientes de la ecuación [1] son nulos.

COMPROBACION

Llevando a la ecuación de la cónica, $X^t M X = O$, la expresión de la columna X de un punto de la recta (en función de los parámetros λ y μ), se tiene:

$$(\lambda\alpha + \mu\beta)^t M(\lambda\alpha + \mu\beta) = O$$

y operando se obtiene rápidamente la ecuación [1] del enunciado (*), que también se puede poner en la forma (si $\mu \neq 0$):

$$(\alpha^t M \alpha)\left(\frac{\lambda}{\mu}\right)^2 + 2(\alpha^t M \beta)\left(\frac{\lambda}{\mu}\right) + (\beta^t M \beta) = O$$

que es de segundo grado, en la incógnita λ/μ. Si las raíces de esta ecuación

(*) Téngase en cuenta que la matriz $\alpha^t M \beta$, como tiene tamaño 1×1, es igual a su traspuesta, la $\beta^t M \alpha$, ya que M es simétrica.

son m_1 y m_2, entonces los puntos de intersección se obtienen tomando $(\lambda, \mu) = (m_1, 1)$ y $(\lambda, \mu) = (m_2, 1)$. Los puntos de intersección son, pues, dos, uno o ninguno según que el discriminante de la última ecuación sea positivo nulo o negativo. En particular, la recta será tangente (hay un solo punto de intersección) cuando se anula el referido discriminante, es decir, cuando se verifica la igualdad [2] del enunciado.

OBSERVACION Y EJEMPLOS

Conviene tener presente que, dadas una recta y una cónica, alguno de sus dos puntos de intersección puede estar en el infinito. Esto acontecerá si, para una solución (λ_1, μ_1) de la ecuación [1]. la primera coordenada (X_0) de $X = \lambda_1 \alpha + \mu_1 \beta$ es cero.

Si, en el proceso anterior, se hubiera recurrido a las coordenadas absolutas (x_1, x_2), en lugar de usar las homogéneas (X_0, X_1, X_2), se habría perdido el rastro de los puntos del infinito, como se puede apreciar en los siguientes ejemplos:

1. Supongamos que se desea hallar la intersección de la recta $y = 3$ y la parábola $y^2 = 6x$. Utilizando las coordenadas absolutas x e y, se tiene un único punto:

$$\left.\begin{matrix} y = 3 \\ y^2 = 6x \end{matrix}\right\}, \quad \left.\begin{matrix} y = 3 \\ 9 = 6x \end{matrix}\right\}, \quad \left.\begin{matrix} y = 3 \\ x = 3/2 \end{matrix}\right\} \qquad (x, y) = (3/2, 3)$$

Ahora bien, operando en coordenadas homogéneas (X_0, X_1, X_2), las cosas quedan así:

$$\left.\begin{matrix} X_2 = 3X_0 \\ X_2^2 = 6X_0X_1 \end{matrix}\right\}, \quad \left.\begin{matrix} X_2 = 3X_0 \\ 9X_0^2 = 6X_0X_1 \end{matrix}\right\}, \quad \left.\begin{matrix} X_2 = 3X_0 \\ X_1 = (3/2)X_0 \text{ ó } X_0 = 0 \end{matrix}\right\}$$

$$(X_0, X_1, X_2) = (\rho, 3/2\rho, 3\rho) \quad \text{ó} \quad (X_0, X_1, X_2) = (0, \rho, 0)$$

El primero de estos puntos es el hallado antes; el segundo es un punto del infinito.

2. Busquemos ahora la intersección de una hipérbola con una de sus asíntotas:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \text{y} \quad \frac{x}{a} = \frac{y}{b}$$

Utilizando coordenadas cartesianas absolutas, nos encontramos con que:

$$\left.\begin{matrix} \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1 \\ \dfrac{x}{a} = \dfrac{y}{b} \end{matrix}\right\}, \quad \left.\begin{matrix} 0 = 1 \\ \dfrac{x}{a} = \dfrac{y}{b} \end{matrix}\right\} \text{ es incompatible}$$

Recurriendo a las coordenadas homogéneas (X_0, X_1, X_2) se tiene:

$$\left.\begin{aligned}\frac{X_1^2}{a^2}-\frac{X_2^2}{b^2}&=X_0^2\\ \frac{X_1}{a}&=\frac{X_2}{b}\end{aligned}\right\}\quad\left.\begin{aligned}0&=X_0^2\\ \frac{X_1}{a}&=\frac{X_2}{b}\end{aligned}\right\}\quad (X_0, X_1, X_2)=(0, \rho a, \rho b)\quad (\rho\neq 0)\quad \text{solución doble}$$

Así pues, esta recta (la asíntota) es tangente a la hipérbola y lo es en el infinito.

EJERCICIO

Hallar las tangentes a la siguiente cónica que son paralelas a la recta $x+y=0$:

$$x^2+2y^2+4xy+4x+2y-5=0$$

RESOLUCION

Las tangentes buscadas tienen ecuación de la forma $x+y=h$; hay que buscar h de manera que la intersección de $x+y=h$ con la cónica sea un punto doble. Llevando, pues, $y=-x+h$ a la ecuación de la cónica, se tiene:

$$x^2+2(-x+h)^2+4x(-x+h)+4x+2(-x+h)-5=0 \quad \text{ó}$$
$$x^2-2x+(2h^2+2h-5)=0$$

Esta ecuación ha de tener una raíz doble, es decir, su discriminante debe ser nulo, o sea:

$$2h^2+2h-4=0, \quad \text{luego} \quad h=1 \quad \text{ó} \quad h=-2$$

Las tangentes pedidas son, pues, las $x+y=1$ y $x+y+2=0$.

CONICAS ORDINARIAS Y DEGENERADAS

[307]

En el plano afín ampliado y respecto de una referencia cartesiana, se considera la cónica cuya matriz es M (simétrica, 3×3). Se dice que la cónica es *ordinaria* o que es *degenerada* según que M sea regular ($\det M\neq 0$) o singular ($\det M=0$)(*), respectivamente.

Si la cónica es degenerada, entonces se llaman *puntos singulares* de la misma a aquellos puntos cuya matriz X (columna de sus coordenadas homogéneas X_0, X_1, X_2) es una solución no nula de la ecuación $MX=O$.

Si P_s es un punto singular de una cónica degenerada, entonces ésta es un par de rectas (distintas, coincidentes o imaginarias) que pasan por P_s.

(*) Esta definición no depende de la referencia cartesiana (es una definición consistente) ya que la anulación de $\det M$ es invariante proyectivo (ver [304]).

DEMOSTRACION

Si P_s es un punto singular de la cónica y P_0 es un punto de la cónica, comprobemos que todos los puntos P de la recta P_sP_0 pertenecen a la cónica. Sean α, β y χ las columnas de coordenadas homogéneas de los puntos P_0, P_s y P (punto genérico de la recta P_sP_0), de manera que $\chi = \lambda\alpha + \mu\beta$, para $(\lambda, \mu) \neq (0, 0)$ recorriendo $\mathbb{R}^2$. Los puntos de intersección de esta recta y la cónica se obtienen, según se vió en [306], [1], para aquellos (λ, μ) que verifican a la ecuación:

$$(\alpha^t M\alpha)\lambda^2 + 2(\alpha^t M\beta)\lambda\mu + (\beta^t M\beta)\mu^2 = O \qquad [1]$$

Como P_0 pertenece a la cónica, es $\alpha^t M\alpha = O$; como P_s es singular, es $M\beta = O$ y, por tanto, $\alpha^t M\beta = O$ y $\beta^t M\beta = O$. Así, pues, todos los coeficientes de la ecuación [1] son nulos y ella se satisface, entonces, para cualesquiera λ y μ, luego todos los puntos P de la recta P_sP_0 son de la cónica.

Resulta, pues, que la cónica degenerada. si tiene algún punto además del P_s, está formada por rectas que pasan por P_s. Más exactamente, la cónica es una pareja de rectas, ya que su ecuación es de segundo grado.

OBSERVACION

Supongamos que se dispone de una cónica degenerada y que P_s es punto singular de ella. Trasladando el origen de coordenadas al punto P_s, la matriz M de la cónica en los nuevos ejes ha de ser tal que

$$M\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix} \quad \text{luego} \quad M = \begin{bmatrix}0 & 0 & 0\\0 & a_{11} & a_{12}\\0 & a_{21} & a_{22}\end{bmatrix}$$

y la ecuación de la cónica (llamando x_1 y x_2 a las coordenadas cartesianas) es:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 0 \qquad [1]$$

Si es $a_{22} \neq 0$, recurriendo a las dos raíces, m_1 y m_2, de la ecuación

$$a_{22}m^2 + 2a_{12}m + a_{11} = 0 \quad (\text{donde } m = x_2/x_1)^{(*)} \qquad [2]$$

la ecuación [1] de la cónica se puede poner en la forma

$$a_{22}(x_2 - m_1x_1)(x_2 - m_2x_1) = 0$$

Por tanto, la cónica es el par de rectas $x_2 = m_1x_1$ y $x_2 = m_2x_1$, que son distintas

(*) Si $a_{22} = 0$, esta ecuación no es de segundo grado. Si $a_{22} = 0$ pero $a_{11} \neq 0$, se puede recurrir a la ecuación $a_{11}n^2 + 2a_{12}n + a_{22} = 0$ (donde $n = x_1/x_2$). Si $a_{22} = a_{11} = 0$, la ecuación dada representa al par de rectas $x_1 = 0$ y $x_2 = 0$ (los dos ejes coordenados).

(reales), coincidentes (real y doble) o imaginarias según que lo sean, respectivamente, las raíces, m_1 y m_2, de la ecuación [2].

EJERCICIO

Comprobar que la siguiente cónica es degenerada y hallar las dos rectas que la forman:

$$2x^2 - y^2 - xy + 3y - 2 = 0$$

RESOLUCION

La cónica tiene por matriz a la

$$M = \begin{bmatrix} -4 & 0 & 3 \\ 0 & 4 & -1 \\ 3 & -1 & -2 \end{bmatrix}$$

Como $\det M = 0$, la cónica es degenerada. Para hallar las dos rectas que constituyen la cónica, basta con considerar a la ecuación de ésta como ecuación de segundo grado en y (x hace papel de parámetro) y despejar esta variable, con lo que se obtiene:

$$y = \frac{3 - x \pm \sqrt{(x-3)^2 + 4(2x^2 - 2)}}{2} = \frac{3 - x \pm \sqrt{(3x-1)^2}}{2} = \begin{cases} x + 1 \\ -2x + 2 \end{cases}$$

Así, pues, las rectas pedidas son

$$x - y + 1 = 0 \quad \text{y} \quad 2x + y - 2 = 0$$

El punto singular de la cónica, por ser el de intersección de estas dos rectas, es el (1/3, 4/3). Para hallar este punto directamente (sin recurrir a las dos rectas anteriores, expresándolo a través de sus coordenadas homogéneas (X_0, X_1, X_2), habría que resolver la ecuación $MX = O$, es decir:

$$\begin{bmatrix} -4 & 0 & 3 \\ 0 & 4 & -1 \\ 3 & -1 & -2 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

cuyas soluciones son:

$$(X_0, X_1, X_2) = (3\rho, \rho, 4\rho)$$

lo que nos conduce al punto de coordenadas absolutas $x = 1/3$, $y = 4/3$, que coincide con el hallado antes.

CONICAS REALES E IMAGINARIAS

[308]

Una cónica que tenga algún punto (uno al menos), se dice que es real; si carece de puntos, se llama imaginaria.

En el plano afín ampliado y respecto de una referencia cartesiana, la cónica cuya matriz es M (simétrica, 3×3) es *imaginaria* o *real* según que, respectivamente, M sea definida o no lo sea, esto es, según que la signatura de M sea $\{3, 0\}$(*) o no lo sea.

NOTA: Si M es definida, entonces la cónica carece de puntos, tanto propios como impropios.

(*) Para señalar que la sig$M$ vale (3, 0) ó (0, 3), se pondrá sig$M = \{3, 0\}$.

DEMOSTRACION

La cónica será imaginaria, es decir, carecerá de puntos, si su ecuación en coordenadas homogéneas, $X^t M X = O$ (donde X es la columna de las coordenadas homogéneas de un punto genérico del plano ampliado), no se verifica para ningún $X \neq O$, lo que significa que M es definida.

EJERCICIO

Averiguar si la siguiente cónica es imaginaria para algunos valores del parámetro $a \in \mathbb{R}$:

$$5x^2 + ay^2 - 8xy + 4x - 2y + 1 = 0$$

RESOLUCION

Diagonalizando por congruencia la matriz M de la cónica, se obtiene fácilmente la matriz D:

$$M = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & -4 \\ -1 & -4 & a \end{bmatrix}; \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a-5 \end{bmatrix}$$

La cónica será imaginaria si los tres elementos de la diagonal de D tienen el mismo signo, esto es, para $a > 5$.

TANGENTES A UNA CONICA

[309]

En el plano afín ampliado y respecto de una referencia cartesiana (sea X la columna de las coordenadas homogéneas de un punto genérico), se considera la cónica ordinaria cuya matriz es M (esto es, cuya ecuación homogénea es $X^tMX = O$); se considera también el punto A cuya columna de coordenadas homogéneas es α. Se verifica que:

1. **Si A pertenece a la cónica, entonces la cónica tiene tangente (única) en A; esta tangente es la recta de ecuación homogénea.**

$$\alpha^tMX = O \qquad [1]$$

(en este caso, a la tangente en A se la llama también polar de A).

2. **Si el punto A no pertenece a la cónica, desde A se pueden trazar dos tangentes (reales o imaginarias) a la cónica; la ecuación conjunta de éstas es:**

$$(\alpha^tMX)^2 = (\alpha^tM\alpha)(X^tMX) \qquad [2]$$

Los dos puntos de tangencia determinan una recta, que se llama *polar* del punto A respecto de la cónica; la ecuación de la polar es $\alpha^tMX = O$.

DEMOSTRACION

Según ya obtuvimos anteriormente (véase [316], [2]) una recta AB es tangente a la cónica si, llamando β a la columna de coordenadas de B, se verifica que:

$$(\alpha^tM\beta)^2 = (\alpha^tM\alpha)(\beta^tM\beta)$$

Si aquí se sustituye la columna fija β por la columna variable X de las coordenadas de un punto genérico del plano, se obtiene, pues, la ecuación del conjunto de las tangentes a la cónica desde A. que es, entonces

$$(\alpha^tMX)^2 = (\alpha^tM\alpha)(X^tMX) \qquad [3]$$

1. Si el punto A pertenece a la cónica, entonces es $\alpha^tM\alpha = O$ y, por ello, la ecuación [3] queda reducida a $(\alpha^tMX)^2 = O$, que no es otra cosa que $\alpha^tMX = O$ (contada dos veces), que es una recta (aquella cuyos coeficientes forman la fila α^tM(*)).
2. Si el punto A no pertenece a la cónica. no se produce la simplificación del caso anterior y la ecuación [3], que es la [2] del enunciado, es de

(*) La fila α^tM no es nula ya que $A(\alpha)$ no es punto singular de la cónica, pues ésta es ordinaria.

segundo grado y, por ello, las rectas tangentes desde A son dos(*). Los puntos de tangencia son los de intersección de éstas (ecuación [2]) con la cónica (ecuación $X^t M X = O$); si para los puntos de tangencia se verifican, entonces, estas dos ecuaciones, también se habrá de verificar la que resulta de imponer en [2] que sea $X^t M X = O$, que es la ecuación $\alpha^t M X = 0$ (dos veces). Por tanto, esta recta contiene a los puntos de tangencia, como había que comprobar.

EJERCICIO

Dada la cónica

$$3x^2 + 2xy - y^2 + 3x - 4y - 5 = 0$$

y el punto $P(a, b)$, hallar: 1) si P es de la cónica, la tangente en P; 2) si P no es de la cónica, la polar de P.

RESOLUCION

Las dos cosas que se piden se obtienen de la misma fórmula: la [1] de [309]; recurriendo a ella, las rectas pedidas son

$$[1 \quad a \quad b]\begin{bmatrix} -5 & 3/2 & -2 \\ 3/2 & 3 & 1 \\ -2 & 1 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ x \\ y \end{bmatrix} = O$$

o sea,

$$3ax + (ay + xb) - by + \tfrac{3}{2}(a + x) - 2(b + y) - 5 = 0$$

o bien

$$(\tfrac{3}{2} + 3a + b)x + (-2 + a - b)y + (-5 + \tfrac{3}{2} - 2) = 0$$

14.2. ANALISIS AFIN DE LAS CONICAS

Aun cuando las cónicas las venimos estudiando, desde el principio, en el plano afín, también es cierto que en lo que se ha dicho en el apartado anterior, que es válido en el espacio afín ampliado, no se ha establecido diferencias entre los puntos propios y los del infinito, por lo que tiene carácter proyectivo (véase [246]). Es a partir de este momento cuando empezaremos a estudiar aquellas propiedades de las cónicas que se originan como consecuencia de dar carácter diferenciado a los puntos del infinito.

(*) Habría que comprobar que estas dos rectas son distintas, cosa que no haremos por ser excesivamente farragoso.

PUNTOS DEL INFINITO; CLASIFICACION AFIN

Según ya se ha hecho notar anteriormente, al considerar las cónicas inmersas en el plano ampliado (con los puntos del infinito), nos vamos a encontrar con que la ecuación de la cónica puede ser satisfecha, además de por sus puntos propios, por algún punto del infinito; de éstos nos ocupamos aquí:

[310]

En el plano afín, se considera la cónica que, respecto de una cierta referencia cartesiana (coordenadas homogéneas X_0, X_1, X_2), admite por ecuación homogénea a:

$$\sum_{i,j=1}^{2} a_{ij}X_iX_j + 2\sum_{h=1}^{2} b_hX_0X_h + cX_0^2 = 0; \quad \text{sea} \quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Se llaman puntos del infinito de la cónica a los puntos del infinito que satisfacen a su ecuación homogénea, esto es, a las soluciones de:

$$\begin{cases} a_{11}X_1^2 + 2a_{12}X_1X_2 + a_{22}X_2^2 = 0^{(*)} \\ X_0 = 0 \end{cases} \qquad [1]$$

Atendiendo a sus puntos del infinito, las cónicas se clasifican como sigue:

La cónica se dice que es:	si en el infinito tiene	
de tipo ELIPTICO	0 puntos (son imaginarios)	$\Leftrightarrow \det A > 0$
de tipo PARABOLICO	1 punto (dos confundidos)	$\Leftrightarrow \det A = 0$ [2]
de tipo HIPERBOLICO	2 puntos (reales y distintos)	$\Leftrightarrow \det A < 0$

(*) Esta ecuación se puede poner (para $X_1 \neq 0$) $a_{22}m^2 + 2a_{12}m + a_{11} = 0$, donde $m = X_2/X_1$; si m_1 y m_2 son las raíces de esta ecuación, los puntos del infinito son los $(0, 1, m_1)$ y $(0, 1, m_2)$ es decir, los del infinito de las rectas de pendientes m_1 y m_2.

COMPROBACION

Los puntos del infinito están caracterizados por tener $X_0 = 0$ (véase [243] y [245]), luego los puntos del infinito de la cónica se obtienen al hacer $X_0 = 0$ en la ecuación de ésta, lo que proporciona el sistema [1]. Este sistema es equivalente (si es $X_1 \neq 0$)(*) al que resulta de dividir por X_1^2, que es:

$$\begin{cases} a_{22}m^2 + 2a_{12}m + a_{11} = 0 \\ X_0 = 0 \end{cases} \qquad \text{donde} \quad m = X_2/X_1$$

(*) La solución $X_1 = 0$ se da si es $a_{22} = 0$; cuando así sea, se cambiarán los papeles de X_1 y X_2; si a_{11} y a_{22} son nulos, la ecuación tiene por soluciones a las $(X_1, X_2) = (0, 1)$ y $(X_1, X_2) = (1, 0)$.

Como el discriminante de esta ecuación de segundo grado es

$$4a_{12}^2 - 4a_{11}a_{22} = -4 \det A$$

resulta que habrá dos, una o ninguna solución según que $\det A$ sea, respectivamente, negativo, nulo o positivo, como había que comprobar.

EJERCICIO

Sabiendo que las dos cónicas siguientes son de tipo parabólico, hallar las relaciones que existen entre α, β y γ:

$$\alpha x^2 + 2\beta xy + \gamma y^2 + 2x + 3y + 1 = 0$$
$$(\alpha + 1)x^2 + 2(\beta + 1)xy + (\gamma + 1)y^2 + 4 = 0$$

RESOLUCION

El que las cónicas sean de tipo parabólico significa que

$$0 = \begin{vmatrix} \alpha & \beta \\ \beta & \gamma \end{vmatrix} = \alpha\gamma - \beta^2 \quad \text{y} \quad 0 = \begin{vmatrix} \alpha + 1 & \beta + 1 \\ \beta + 1 & \gamma + 1 \end{vmatrix} = (\alpha\gamma - \beta^2) + (\alpha + \gamma - 2\beta)$$

Este sistema de ecuaciones conduce obviamente a $\alpha = \beta = \gamma$.

ELIPSES, PARABOLAS E HIPERBOLAS

[311]

Una cónica no degenerada, atendiendo a sus puntos del infinito, se llama:

- *Elipse* si, además de ser una cónica ordinaria, es de tipo elíptico (o sea, sin puntos en el infinito).
- *Parábola* si, además de ser una cónica ordinaria, es de tipo parabólico (o sea, tangente al infinito).
- *Hipérbola* si, además de ser una cónica ordinaria, es de tipo hiperbólico (o sea, con dos puntos en el infinito).

La cónica que, respecto de una referencia cartesiana del plano, tiene a M por matriz (simétrica; 3×3) y a A como submatriz de los términos cuadráticos, es elipse, parábola o hipérbola si se verifica que $\det M \neq 0$ y $\det A$ es, respectivamente, positivo, nulo o negativo.

COMPROBACION

El resultado que se acaba de enunciar es consecuencia obvia de [307] (la cónica es ordinaria si $\det M \neq 0$) y de [310] (clasificación afín).

EJERCICIO

Compruébese que las siguientes cónicas C_1, C_2 y C_3 son, respectivamente, una elipse, una parábola y una hipérbola:

$$C_1: x^2 + 2xy + 6y^2 + 8y + 2 = 0$$
$$C_2: x^2 - 2xy + y^2 + 6x + 1 = 0$$
$$C_3: 2x^2 - xy - y^2 + 2x + y + 3 = 0$$

RESOLUCION

Para cada una de las cónicas dadas, recurriendo a su matriz M y a la submatriz A de sus términos cuadráticos, según lo dicho en [311], se obtiene:

- Para C_1: $\det A = 2 > 0$ y $\det M = -12 \neq 0$, luego C_1 es una elipse.
- Para C_2: $\det A = 0$ y $\det M = -9 \neq 0$, luego C_2 es una parábola.
- Para C_3: $\det A = -9/4$ y $\det M = -27/4 \neq 0$, luego C_3 es una hipérbola.

EJERCICIO

Averiguar de qué tipo es la siguiente cónica, en función del valor que tome el parámetro $a \in \mathbb{R}$:

$$x^2 + y^2 + 2axy + 2x + 1 = 0$$

RESOLUCION

Los determinantes de la matriz M de la cónica y de su submatriz A de términos cuadráticos valen:

$$\det M = -a^2 \quad ; \quad \det A = 1 - a^2$$

Por tanto, en función de los valores de a, se verifica que:

- Si $a = 0$, la cónica es $y^2 + (x + 1)^2 = 0$ y consta de un sólo punto: el $(x, y) = (1, 0)$. Puede decirse que se trata de un par de rectas imaginarias que pasan por dicho punto, pues se puede poner como

$$[y + i(x + 1)][y - i(x + 1)] = 0$$

- Si $0 < |a| < 1$, la cónica es una elipse, pues $\det M \neq 0$ y $\det A > 0$.
- Si $a = \pm 1$, la cónica es una parábola, pues $\det M \neq 0$ y $\det A = 0$.
- Si $|a| > 1$, la cónica es una hipérbola, pues $\det M \neq 0$ y $\det A < 0$.

CLASIFICACION GENERAL DE LAS CONICAS

[312]

En el siguiente cuadro se plasma la clasificación general de las cónicas; en él, M es la matriz de la cónica y A es su submatriz de términos cuadráticos (en una referencia cartesiana cualquiera):

$\det A > 0$ CONICA DE TIPO ELIPTICO	$\operatorname{Sig} M = \{3, 0\}$ (1)	ELIPSE IMAGINARIA
	$\operatorname{Sig} M = \{2, 1\}$ (1)	ELIPSE REAL
	$\det M = 0$	UN PUNTO (dos rectas imaginarias que se cortan en el punto)
$\det A < 0$ CONICA DE TIPO HIPERBOLICO	$\det M \neq 0$	HIPERBOLA
	$\det M = 0$	DOS RECTAS (REALES) CONCURRENTES
$\det A = 0$ CONICA DE TIPO PARABOLICO	$\det M \neq 0$	PARABOLA
	$\det M = 0$	DOS RECTAS PARALELAS (distintas, coincidentes o imaginarias) (2)

(1) $Sig\, M = \{p, q\}$ significa que la signatura de M vale (p, q) ó (q, p).
(2) Las rectas paralelas son imaginarias, distintas o coincidentes, según que, respectivamente, sea $\operatorname{sig} M = \{2, 0\}$, $\operatorname{sig} M = \{1, 1\}$, $\operatorname{rang} M = 1$.

COMPROBACION

Salvo alguna cuestión de detalle, no hay nada realmente novedoso. En el cuadro del enunciado se recogen las clasificaciones anteriores; esto es: en ordinarias ($\det M \neq 0$) y degeneradas ($\det M = 0$), que vimos en [307]; en reales ($\operatorname{sig} M \neq \{3, 0\}$) e imaginarias ($\operatorname{sig} M = \{3, 0\}$), que vimos en [308]; la clasificación afín (en función del signo de $\det A$), que vimos en [310]; la definición de elipse, parábola e hipérbola, que vimos en [311]. Solamente hay que hacer notar que una cónica degenerada (par de rectas) de tipo elíptico (sin puntos en el infinito), de tipo hiperbólico (con dos puntos en el infinito) o de tipo parabólico (con un punto en el infinito) ha de ser forzosamente un par de rectas imaginarias, un par de rectas reales concurrentes o un par de rectas paralelas, respectivamente.

EJERCICIO

Clasificar, en función del valor que tome el parámetro $a \in \mathbb{R}$, la siguiente cónica:

$$x^2 + 2ay^2 - 2axy - 2x + 4ay = 0$$

RESOLUCION

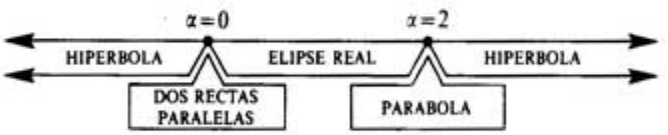
Diagonalizando por congruencia la matriz M de esta cónica se llega fácilmente a la siguiente matriz D:

$$M = \begin{bmatrix} 0 & -1 & 2a \\ -1 & 1 & -a \\ 2a & -a & 2a \end{bmatrix}; \quad D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2a \end{bmatrix}$$

Como, además, $\det M = -2a$ y $\det A = a(2 - a)$ (donde A es la matriz de los términos cuadráticos de la cónica), resulta que:

- Si $a < 0$, es $\det A < 0$ y $\det M \neq 0 \rightarrow$ hipérbola.
- Si $a = 0$, es $\det A = 0$ y $\det M = 0 \rightarrow$ rectas paralelas.
- Si $0 < a < 2$, es $\det A > 0$ y $\text{sig } M = \{2, 1\} \rightarrow$ elipse real.
- Si $a = 2$, es $\det A = 0$ y $\det M \neq 0 \rightarrow$ parábola.
- Si $2 < a$, es $\det A < 0$ y $\det M \neq 0 \rightarrow$ hipérbola.

Este resultado puede representarse recurriendo al siguiente esquema:



OBSERVACION

El cuadro de [312], de la clasificación general. puede ponerse, también, en la forma:

$\text{Sig } M = \{3, 0\}$	y	$\text{Sig } A = \{2, 0\}$	$\leftrightarrow$ Elipse imaginaria
$\text{Sig } M = \{2, 1\}$	y	$\text{Sig } A = \{2, 0\}$	$\leftrightarrow$ Elipse real
$\text{Sig } M = \{2, 1\}$	y	$\text{Sig } A = \{1, 1\}$	$\leftrightarrow$ Hipérbola
$\text{Sig } M = \{2, 1\}$	y	$\text{Sig } A = \{1, 0\}$	$\leftrightarrow$ Parábola

Para comprobar que lo aquí expresado equivale a lo dicho en [312], téngase en cuenta las siguientes consideraciones. Como A es de tamaño 2×2, la condición $\det A > 0$ equivale a $\text{sig } A = \{2, 0\}$ y la condición $\det A < 0$ equivale a que sea $\text{sig } A = \{1, 1\}$. Como $A \neq O$, la condición $\det A = 0$ equivale a $\text{sig } A = \{1, 0\}$. Salvo para las cónicas imaginarias, en las que $\text{sig } M = \{3, 0\}$, la condición $\det M \neq 0$ equivale (por ser M de tamaño 3×3) a $\text{sig } M = \{2, 1\}$.

CONICAS CON CENTRO

[313]

En el plano afín y respecto de una cierta referencia cartesiana (coordenadas x_1, x_2), se considera la cónica que admite por ecuación a:

$$\sum_{i,j=1}^{2} a_{ij}x_ix_j + 2\sum_{h=1}^{2} b_hx_h + c = 0 \quad \text{cuya matriz es} \quad M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix}$$

Se dice que un punto P_c es *centro* de la cónica si es centro de simetría de la misma. A este respecto, se verifica que:

I. Si un punto $P_c(x_1, x_2)$ es solución del sistema de ecuaciones:

$$\sum_{j=1}^{2} a_{ij}x_j = -b_i \quad (\text{para } i = 1, 2) \quad \text{o} \quad AX = -B \qquad [1]$$

entonces P_c es centro de la cónica. En relación con el centro, de la discusión de [1], se obtiene:

1. Si $\delta_0 = \det A \neq 0$, entonces la cónica tiene un centro, que es:

$$P_c\left(\frac{\delta_1}{\delta_0}, \frac{\delta_2}{\delta_0}\right) \qquad [2]$$

donde, δ_0, δ_1 y δ_2 son los adjuntos de los elementos de la primera fila de M. Las elipses, las hipérbolas y los pares de rectas concurrentes tienen centro (para estas últimas, el centro es su punto singular).

2. Las parábolas (que tienen $\delta_0 = \det A = 0$) carecen de centro(*).

II. Si una cónica tiene centro, al que llamaremos $P_c(\alpha_1, \alpha_2)$, y se traslada la referencia cartesiana para que tenga su origen en P_c (nuevas coordenadas $x_1' = x_1 - \alpha_1$ y $x_2' = x_2 - \alpha_2$), entonces la ecuación de la cónica toma la siguiente forma (que carece de términos lineales):

$$\sum_{i,j=1}^{2} a_{ij}x_i'x_j' + c' = 0, \quad \text{donde} \quad c' = c + \sum_{h=1}^{2} b_h\alpha_h \qquad [3]$$

(*) Un par de rectas paralelas (det $A = 0$, det $M = 0$) tiene infinitos centros, que son soluciones de [1]. Este sistema tiene, ahora, infinitas soluciones pues rang A = rang $AB < 2$.

DEMOSTRACION

I. Si $P_c(\alpha_1, \alpha_2)$ es un punto tal que, al tomarle como origen de la referencia (trasladando los ejes; nuevas coordenadas $x_i' = x_i - \alpha_i$), la nueva ecuación

de la cónica no tuviera términos lineales, entonces P_c sería centro de simetría de la cónica, ya que si un punto $P(x_1', x_2')$ verifica a dicha ecuación, entonces es evidente que también la verificaría su simétrico $P'(-x_1', -x_2')$ respecto de P_c. Vamos, pues, a buscar $P_c(\alpha_1, \alpha_2)$ de manera que, con el cambio de coordenadas $x_1' = x_1 - \alpha_1$ y $x_2' = x_2 - \alpha_2$, se anulen los términos lineales. Al hacer dicho cambio, la ecuación de la cónica queda:

$$\sum_{i,j=1}^{2} a_{ij}x_i'x_j' + 2\sum_{i=1}^{2}\left(\sum_{j=1}^{2} a_{ij}\alpha_i + b_i\right)x_i' + \\ + \left(\sum_{i,j=1}^{2} a_{ij}\alpha_i\alpha_j + 2\sum_{k=1}^{2} b_k\alpha_k + c\right) = 0 \qquad [A]$$

y por tanto, la anulación de los términos lineales se produce si

$$\sum_{j=1}^{n} a_{ij}\alpha_j = -b_i \quad (\text{para } i = Z1, 2) \qquad [B]$$

que es el primer resultado, [1], del enunciado.

1. Si $\delta_0 = \det A \neq 0$, el sistema de ecuaciones [1] es de Cramer, luego tiene solución única. Esta solución es la dada en [2], esto es, $x_i = \delta_i/\delta_0$ (para $i = 1, 2$), ya que:

$$\sum_{j=1}^{n} a_{ij}\left(\frac{\delta_i}{\delta_0}\right) + b_i = \frac{1}{\delta_0}(b_i\delta_0 + a_{i1}\delta_1 + a_{i2}\delta_2) = \frac{1}{\delta_0}\,0 = 0$$

puesto que el último paréntesis es la suma de los productos de los elementos de la fila $i + 1$ de M (esto es, los b_i, a_{i1}, a_{i2}) por los adjuntos de la primera fila de M, que vale 0.

Obsérvese, además, que si la cónica es degenerada, o sea, si $\det M = 0$, entonces P_c es punto singular de la cónica, ya que

$$M\begin{bmatrix}\delta_0\\ \delta_1\\ \delta_2\end{bmatrix} = \begin{bmatrix}\det M\\ 0\\ 0\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

2. Si $\delta_0 = \det A = 0$, entonces la matriz del sistema de ecuaciones [1] es singular y, según el teorema de Rouché, para que dicho sistema tenga solución (que no será única) es necesario y suficiente que sea $\text{rang}\,A = \text{rang}\,\mathcal{AB}$. Para que se cumpla esta condición, M ha de ser singular (pues $\text{rang}\,A < 2$ y $\text{rang}\,A = \text{rang}\,\mathcal{AB}$ conducen a $\text{rang}\,M < 3$) y, por ello, la referida condición no se puede cumplir para las parábolas, que carecen, pues, de centro.

II. Como $P_c(\alpha_1, \alpha_2)$ verifica a la ecuación [1], resulta que, para P_c, la ecuación

[A] tiene nulos los términos lineales y es del tipo [3] que se dice en el enunciado, ya que, además:

$$c' = \sum_{i=1}^{2}\left(\sum_{j=1}^{2} a_{ij}\alpha_j\right)\alpha_i + 2\sum_{h=1}^{2} b_h\alpha_h + c = \sum_{i=1}^{2} - b_i\alpha_i + 2\sum_{h=1}^{2} b_h\alpha_h + c =$$

$$= \sum_{h=1}^{2} b_h\alpha_h + c$$

EJERCICIO

Hallar los centros (cuando existan) de las siguientes cónicas, para $a \in \mathbb{R}$:

$$x^2 + y^2 + 2xy + 2x + 2y + 2 = 0$$

RESOLUCION

La ecuación $AX = -B$, que proporciona los centros (x, y), es (véase [313], [1]):

$$\left.\begin{matrix} 1 + x + ay = 0 \\ 1 + \alpha x + y = 0 \end{matrix}\right\} \quad A = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}; \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

La discusión de este sistema conduce fácilmente a:

- Si $a \neq \pm 1$, el centro es $(-1/(a+1),\ -1/(a+1))$; la cónica es elipse o hipérbola.
- Si $a = -1$, no hay centro; la cónica es una parábola.
- Si $a = 1$, hay una recta de centros (la $x + y + 1 = 0$); la cónica es un par de rectas paralelas (imaginarias).

DIRECCION DE UNA PARABOLA

[314]

En el plano afín y respecto de una cierta referencia cartesiana, se considera la parábola cuya matriz es M; sea A su submatriz de términos cuadráticos. Llamemos δ_0, δ_1 y δ_2 a los adjuntos de los elementos de la primera fila de M (nótese que $\delta_0 = \det A = 0$, pues se habla de una parábola). Entonces se verifica que:

1. El punto del infinito de la parábola, P_∞, tiene por coordenadas homogéneas a las $P_\infty(0, \delta_1, \delta_2)$; esto es, P_∞ es el punto del infinito de la dirección del vector $\bar{\delta}(\delta_1, \delta_2)$, a la que se llama *dirección de la parábola*.
2. El vector $\bar{\delta}(\delta_1, \delta_2)$ es un vector propio de la matriz A, correspondiente a su autovalor $\lambda = 0$.

NOTA: Todo lo anterior es válido, también, si la cónica es un par de rectas paralelas. En este caso, la dirección de $\bar{\delta}$ es la de las dos rectas.

DEMOSTRACION

Nótese primeramente que, como $\det A = 0$, la matriz A admite el autovalor $\lambda = 0$. Obsérvese también que, llamando

$$\Delta = \begin{bmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \delta_1 \\ \delta_2 \end{bmatrix} \quad \text{es} \quad M\Delta = \begin{bmatrix} \det M \\ 0 \\ 0 \end{bmatrix} \qquad [1]$$

ya que la suma de los productos de los elementos δ_0, δ_1, δ_2 (que son los adjuntos de los elementos de la primera fila de M), por los respectivos elementos de la fila i-ésima de M vale: $\det M$ si $i = 1$ y 0 si $i \neq 1$.

1. De [1] se deduce obviamente que $\Delta^t M \Delta = O$, lo que significa que el punto $P_\infty(0, \delta_1, \delta_2)$ pertenece a la cónica.
2. Expresando la matriz M en bloques (en función de A, matriz de términos cuadráticos, de B, matriz de términos lineales, y de c, término independiente), la igualdad [1] se puede poner en la forma

$$\begin{bmatrix} \det M \\ 0 \\ 0 \end{bmatrix} = M\Delta = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \begin{bmatrix} 0 \\ \delta_1 \\ \delta_2 \end{bmatrix}$$

De esta igualdad, multiplicando por bloques, se desprende que

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = A \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}$$

es decir, que el vector $\bar{\delta}\,(\delta_1, \delta_2)$ es vector propio de A correspondiente a su autovalor $\lambda = 0$.

ASINTOTAS DE UNA HIPERBOLA

[315]

En el plano afín y respecto de una referencia cartesiana (coordenadas x, y), se considera la hipérbola que admite por ecuación a:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c = 0$$

Las dos raíces, m_1 y m_2, de la ecuación $a_{22}m^2 + 2a_{12}m + a_{11} = 0$ [1] son las pendientes ($m = y/x$) de los dos puntos del infinito de la hipérbola [2]. Las tangentes a la hipérbola en sus puntos del infinito son las rectas de pendientes m_1 y m_2 que pasan por su centro; estas rectas se llaman *asíntotas* de la hipérbola.

---

(1) Las raíces de esta ecuación son reales pues su discriminante es $a_{12}^2 - a_{11}a_{22} > 0$, ya que la cónica es hipérbola. Si fuese $a_{22} = 0$, se cambian los papeles de x e y.

(2) Estos puntos del infinito tienen por coordenadas a $(0, 1, m_1)$ y $(0, 1, m_2)$.

COMPROBACION

Los puntos del infinito de la hipérbola son los puntos de coordenadas homogéneas (0, 1, m) que verifican a su ecuación (se llaman aquí t, x e y a las coordenadas homogéneas); esta es:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + (2b_1x + 2b_2y)t + ct^2 = 0$$

Haciendo aquí $t = 0$, $x = 1$ e $y = m$, se obtiene la ecuación del enunciado:

$$a_{22}m^2 + 2a_{12}m + a_{11} = 0$$

Las rectas cuya pendiente son m_1 y m_2, como pasan por uno de los puntos del inifnito de la hipérbola, la cortarán en un segundo punto (sólo uno); cuando este punto sea también el del infinito, la recta en cuestión será tangente a la hipérbola en el infinito, es decir, será una de sus asíntotas. Las rectas de pendientes m_1 y m_2 que pasan por el centro de la hipérbola son sus asíntotas ya que, de no ser así, cortarían a la hipérbola en un punto propio y, como el centro es centro de simetría de la hipérbola, también la cortarían en el punto simétrico de aquél, luego la recta y la hipérbola tendrían tres puntos en común (el del infinito y los otros dos), lo cual no es posible.

14.3. ANALISIS METRICO DE LAS CONICAS

En nuestro estudio de las cónicas, han quedado para el final, como no podía menos, todo lo relacionado con los ángulos y las distancias. Para no complicar las cosas más allá de lo necesario, a partir de este momento los sistemas de referencia que manejemos serán rectangulares, es decir, de los que tienen la base de vectores ortonormal.

CAMBIO DE COORDENADAS RECTANGULARES

[316]

En el plano afín euclídeo, se considera una cónica cualquiera. Supongamos que, en una cierta referencia rectangular, M es la matriz de la cónica y que A es su submatriz de los términos cuadráticos. Supongamos también que, en otra referencia rectangular, las referidas matrices son M' y A'.

Si Q^* es la matriz del correspondiente cambio de coordenadas (véase [303]), esto es, si:

$$Q^* = \begin{bmatrix} 1 & O \\ C & Q \end{bmatrix}, \text{ donde, } \begin{cases} Q = \text{matriz ortogonal correspondiente al cambio de la base de la referencia rectangular} \\ C = \text{columna de coordenadas del nuevo origen} \end{cases}$$

$$\text{entonces: } \begin{cases} M' = Q^{*t}MQ^* & (M \text{ y } M' \text{ son matrices congruentes}) \\ A' = Q^tAQ & (A \text{ y } A' \text{ son ortogonalmente congruentes}) \end{cases}$$

COMPROBACION

Cuanto acabamos de decir, salvo lo que se derive de que Q es ahora ortogonal, ya está comprobado en [303]. La única novedad estriba ahora en que, por ser Q ortogonal (cosa que allí no ocurría), las matrices A y A' son, no sólo congruentes, sino ortogonalmente congruentes ($A' = Q^t A Q$, siendo Q ortogonal) o, si se prefiere, ortogonalmente semejante, que es la misma cosa.

EJERCICIO

En el plano afín euclídeo y respecto de unos ejes cartesianos rectangulares (coordenadas x_1, x_2), se considera la cónica:

$$\sum_{i,j=1}^{2} a_{ij} x_i x_j + 2 \sum_{h=1}^{2} b_h x_h + c = 0$$

Girar los ejes dados un ángulo θ, que se pide, de manera que en las nuevas coordenadas (x'_1, x'_2), la ecuación de la cónica no tenga términos en $x_1 x_2$:

$$a'_{11} x'^2_1 + a'_{22} x'^2_2 + 2 \sum_{h=1}^{2} b'_h x'_h + c' = 0 \quad (a'_{12} = a'_{21} = 0)$$

o sea, tal que la nueva matriz de términos cuadráticos, $A' = [a'_{ij}]$, sea diagonal.

RESOLUCION

De acuerdo con [313], la matriz A' y, en particular el elemento a'_{12} valdrán:

$$A' = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

$$a'_{12} = \operatorname{sen}\theta \cos\theta (a_{22} - a_{11}) + a_{12}(\cos^2\theta - \operatorname{sen}^2\theta) =$$

$$= \tfrac{1}{2}(a_{22} - a_{11}) \operatorname{sen} 2\theta + a_{12} \cos 2\theta$$

Por tanto, a'_{12} será nulo si

$$\operatorname{tg} 2\theta = \frac{2a_{12}}{a_{11} - a_{22}}$$

Nótese que la solución que se ha encontrado no es válida si $a_{11} = a_{22}$. Cuando sea $a_{11} = a_{22}$, la anterior ecuación da $a'_{12} = a_{12} \cos 2\theta$; luego ahora la anulación de a'_{12} se consigue tomando $\theta = \pi/4$, con lo que (llamando $a_{11} = a_{22} = a$ y $a_{12} = b$):

$$A' = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} a & b \\ b & a \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ 0 & a-b \end{bmatrix}$$

OBSERVACION

Sea M la matriz de una cónica, en una determinada referencia cartesiana del plano. Supongamos que se cambia de referencia y que M' es matriz de la cónica en la nueva referencia, con lo que también son matrices de la cónica, en esta nueva referencia, todas las $\rho M'$ para $\rho \neq 0$. Así, pues, la relación $M' = Q^{*t} M Q^*$, que se obtuvo en [316], deberá sustituirse por la

$$M' = Q^{*t}(\rho M) Q^* \quad \text{(para algún } \rho \neq 0)$$

INVARIANTES METRICOS

[317]

En el plano afín euclídeo y respecto de una cierta referencia rectangular, se considera la cónica cuya matriz es M; sea $A = [a_{ij}]$ su submatriz de los términos cuadráticos y llamemos λ_1 y λ_2 a los autovalores de A. Supongamos que, en una nueva referencia rectangular, los anteriores elementos son ahora M', $A' = [a'_{ij}]$, λ'_1 y λ'_2, respectivamente. Se verifica, entonces, que existe cierto $\rho \in \mathbb{R}$ tal que:

1. $\det M' = \rho^3 (\det M)$
2. $\det A' = \rho^2 (\det A)$
3. $a'_{11} + a'_{22} = \rho(a_{11} + a_{22})$
4. $\lambda'_1 = \rho\lambda_1$
5. $\lambda'_2 = \rho\lambda_2$

(El par de relaciones 2 y 3 es equivalente al par 4 y 5)

Por ello, para señalar este resultado, se dice que $\det M$, $\det A$, $a_{11} + a_{22}$, λ_1 y λ_2 son invariantes métricos de las cónicas.

DEMOSTRACION

Sea Q^* la matriz de cambio de referencia (véase [316]) que se descompone en los siguientes bloques:

$$Q = \begin{bmatrix} 1 & O \\ C & Q \end{bmatrix}, \qquad \text{donde} \begin{cases} C = \text{columna de coordenadas del nuevo origen} \\ Q = \text{matriz ortogonal asociada al cambio de base} \end{cases}$$

Según ya sabemos (de [303]), en la nueva referencia la cónica admite por matriz a $Q^{*t} M Q^*$ y la correspondiente submatriz de los términos cuadráticos es la $Q^t A Q$. Por tanto, como todas las matrices de una variedad cuadrática (en una misma referencia) son proporcionales, existe algún escalar $\rho \in \mathbb{R}$, $\rho \neq 0$, tal que

$$M' = \rho Q^{*t} M Q^* \quad \text{y} \quad A' = \rho Q^t A Q \qquad [1]$$

Dado que $\det Q^* = \det Q = \pm 1$, pues Q es ortogonal, y como M y M' tienen tamaño 3×3, tomando determinantes en la primera de las igualdades [1], se obtiene:

$$\det M' = \rho^3 (\det Q^*)^2 \det M = \rho^3 (\pm 1)^2 \det M = \rho^3 \det M$$

Por otra parte, como A es una matriz simétrica. A es ortogonalmente diagonalizable (véase [178], 2), esto es, existe una matriz ortogonal P, con $P^t = P^{-1}$, tal que $A = P^{-1}DP = P^tDP$, donde D es la matriz diagonal de los autovalores λ_1 y λ_2 de A (los autovalores son los elementos de la diagonal de D). De la segunda de las igualdades [1] se deduce que

$$A' = \rho Q^t A Q = \rho Q^t P^t D P Q = (PQ)^t \rho D (PQ)$$

Como PQ es ortogonal, pues lo son P y Q, la última igualdad nos permite asegurar que la matriz diagonal ρD es la que tiene en su diagonal a los autovalores de A', esto es, que $\lambda_i' = \rho\lambda_i$ (para $i = 1, 2$), como había que comprobar.

Obsérvese finalmente que las relaciones 2 y 3 equivalen a las 4 y 5. Así es, ya que, como [véase [164]):

$$\begin{cases} \lambda_1' \text{ y } \lambda_2' \text{ son las raíces de } \lambda^2 - (a_{11}' + a_{22}')\lambda + \det A' = 0 \\ \rho\lambda_1 \text{ y } \rho\lambda_2 \text{ son las raíces de } \lambda^2 - (\rho a_{11} + \rho a_{22})\lambda + \det(\rho A) = 0 \end{cases}$$

las raíces de una y otra ecuación serán iguales (como dicen las relaciones 4 y 5) si y sólo si sus respectivos coeficientes son proporcionales (como dicen las relaciones 2 y 3).

EJERCICIO

Se sabe que una misma cónica tiene, en dos sistemas de referencia cartesianos rectangulares diferentes, ecuaciones del tipo:

$$\begin{aligned} &\text{en la primera referencia:} && x^2 + ay^2 = 1 \\ &\text{en la segunda referencia:} && x'y' = k \end{aligned}$$

Hallar a y k.

RESOLUCION

Recurriendo a las invariantes métricas, se obtiene que, para cierto $\rho \in \mathbb{R}$, ha de ser:

$$\det M' = \rho^3 \det M \quad \rightarrow \quad k/4 = \rho^3(-a)$$

$$\det A' = \rho^2 \det A \quad \rightarrow -1/4 = \rho^2(a)$$

$$a_{11}' + a_{22}' = \rho(a_{11} + a_{22}) \rightarrow \quad 0 = \rho(1 + a)$$

luego $a = -1$ y $k = \pm 1/2$.

ECUACION REDUCIDA DE UNA CONICA CON CENTRO(*)

[318]

En el plano afín euclídeo y respecto de una referencia rectangular, se considera la cónica con centro cuya matriz es M; sea $A = [a_{ij}]$ su submatriz de los términos cuadráticos (nótese que $\det A \neq 0$, pues la cónica tiene centro). Se llaman *ejes de la cónica* a las dos rectas (perpendiculares entre sí) que pasan por su centro y tienen las direcciones de los vectores de una base ortonormal de autovectores de la matriz A. Las pendientes de estas rectas son las raíces de la ecuación $a_{12}m^2 + (a_{11} - a_{22})m - a_{12} = 0$.

La ecuación de la cónica (con centro) respecto de sus ejes (referencia rectangular; coordenadas x, y) se llama *ecuación reducida* de la cónica y es la siguiente:

$$\lambda_1 x^2 + \lambda_2 y^2 + \frac{\det M}{\det A} = 0 \quad (\lambda_1 \text{ y } \lambda_2 \text{ autovalores de } A) \qquad [1]$$

Los ejes de una cónica (con centro) son ejes de simetría de la misma.

NOTA: En lugar de recurrir a la expresión [1], se puede poner que la ecuación de la cónica es del tipo $\alpha x^2 + \beta y^2 = \gamma$ y hallar α, β y γ expresando que $\det M$, $\det A$ y $a_{11} + a_{22}$ son invariantes, esto es, que se verifica que $\det M = -\alpha\beta\gamma$, $\det A = \alpha\beta$ y $a_{11} + a_{22} = \alpha + \beta$.

DEMOSTRACION

Antes que nada, recordemos que la matriz A, por ser simétrica, es ortogonalmente diagonalizable (véase [178], 2), es decir, existe una matriz ortogonal Q tal que $Q^t A Q = D$ es la matriz diagonal de los autovalores de A (aquella cuya diagonal está formada por los autovalores λ_1 y λ_2 de A); las columnas de Q son vectores propios, $\bar{u}_1$ y $\bar{u}_2$, de A correspondientes a sus autovalores λ_1 y λ_2, que forman una base ortonormal.

Para empezar, traslademos los ejes de la referencia y pongámoslos con origen en el centro (nuevas coordenadas x'_1, x'_2); en estos nuevos ejes, la cónica tiene (véase [313], II) ecuación del tipo.

$$\sum_{i,j=1}^{2} a_{ij} x'_i x'_j + \gamma = 0 \quad (\text{para algún } \gamma \in \mathbb{R})$$

Cambiemos ahora los vectores de la base (de la referencia cartesiana rectangular dada) por los vectores propios $\bar{u}_1$ y $\bar{u}_2$, de los que hablábamos al principio (columnas de la matriz Q); en la nueva referencia que así se obtiene (ahora no

---

(*) Elipses, hipérbolas y pares de rectas concurrentes.

se cambia el origen; nuevas coordenadas x, y). la matriz M' de la cónica será (véase [316]):

$$M' = \begin{bmatrix} 1 & O \\ O & Q^t \end{bmatrix} \begin{bmatrix} \gamma & O \\ O & A \end{bmatrix} \begin{bmatrix} 1 & O \\ O & Q \end{bmatrix} = \begin{bmatrix} \gamma & O \\ O & Q^t A Q \end{bmatrix} = \begin{bmatrix} \gamma & O \\ O & D \end{bmatrix}$$

es decir, en esta referencia la ecuación de la cónica es

$$\lambda_1 x^2 + \lambda_2 y^2 + \gamma = 0$$

Recurriendo ahora a los invariantes métricos (véase [317]), resulta que (el ρ que figura en [325] vale $\rho = 1$, pues A y A' tienen los mismos autovalores):

$$\left.\begin{aligned} \det M = \det M' = \lambda_1 \lambda_2 \gamma \\ \det A = \det A' = \lambda_1 \lambda_2 \end{aligned}\right\} \quad \text{luego} \quad \gamma = \frac{\det M}{\det A}$$

que llevado a [A] nos proporciona la ecuación reducida [1] que buscamos. A la vista de esta ecuación, es obvio que la cónica es simétrica respecto de sus ejes.

Si m es la pendiente de uno cualquiera de los ejes, entonces el vector $(1, m)$ es vector propio de A correspondiente a uno de sus autovalores, que llamaremos λ. Así, pues, se verifica que:

$$\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \begin{bmatrix} 1 \\ m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{o sea} \quad \begin{cases} (a_{11} - \lambda) + m a_{12} = 0 \\ a_{21} + (a_{22} - \lambda) m = 0 \end{cases}$$

Eliminando λ entre estas dos últimas ecuaciones, se obtiene

$$a_{12} m^2 + (a_{11} - a_{12}) m - a_{12} = 0$$

que es, entonces, la ecuación cuyas raíces son las pendientes de los ejes de la cónica.

EJERCICIO

Hallar la ecuación reducida de la cónica

$$3x^2 + 2xy + 3y^2 - 2x - 4y + 1 = 0$$

RESOLUCION

Las matrices de la cónica y de sus términos cuadráticos son

$$M = \begin{bmatrix} 1 & -1 & -2 \\ -1 & 3 & 1 \\ -2 & 1 & 3 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

Como $\det M = -3 \neq 0$ y $\det A = 8 > 0$, la cónica es una elipse (real), cuya ecuación reducida es del tipo $\alpha x^2 + \beta y^2 = 1$. En lugar de aplicar, sin más, la

fórmula [1] de [318], vamos a obtener dicha ecuación reducida recurriendo a los invariantes: como para la nueva ecuación es $\det M' = -\alpha\beta$, $\det A' = \alpha\beta$ y $a'_{11} + a'_{22} = \alpha + \beta$, ha de existir un $\rho \neq 0$ tal que

$$-\alpha\beta = \rho^3(-3) \quad ; \quad \alpha\beta = \rho^2(8) \quad ; \quad \alpha + \beta = \rho(6)$$

De estas ecuaciones se obtiene que $\rho = 8/3$, $\alpha = 32/3$ y $\beta = 16/3$, luego la ecuación reducida es:

$$\frac{32}{3}x^2 + \frac{16}{3}y^2 = 1$$

ECUACION REDUCIDA DE UNA PARABOLA

[319]

En el plano afín euclídeo y respecto de una referencia rectangular, se considera la parábola cuya matriz es M; sea $A = [a_{ij}]$ su submatriz de los términos cuadráticos. Llamemos δ_0, δ_1 y δ_2 a los adjuntos de los elementos de la primera fila de M (nótese que $\delta_0 = \det A = 0$, pues se habla de una parábola); el vector $\bar{\delta} = (\delta_1, \delta_2)$ tiene la dirección de la parábola.

En la parábola hay un punto, llamado *vértice*, en el que la tangente es perpendicular a $\bar{\delta}$; la recta que pasa por el vértice y tiene la dirección de $\bar{\delta}$ se llama *eje de la parábola* (es eje de simetría de la misma). El eje es la polar (véase [309], 2) del punto del infinito de la dirección perpendicular a $\bar{\delta}$(*), es decir, es la recta de ecuación $[0\, a_{11}\, a_{22}]MX = O$ (donde X es la columna de las coordenadas homogéneas de un punto genérico del plano).

La ecuación de la parábola respecto de su eje («eje de la x») y de su tangente en el vértice («eje de la y») se llama *ecuación reducida* y es la:

$$y^2 = 2px \quad \text{donde} \quad p = \sqrt{-\frac{\det M}{(a_{11} + a_{22})^3}}$$

(*) Las direcciones del eje y de la tangente en el vértice son las de los subespacios propios de la matriz A, correspondientes a sus autovalores $\lambda = 0$ y $\lambda \neq 0$, respectivamente; esto es, los de los vectores $\bar{\delta} = (\delta_1, \delta_2)$ y $\bar{\delta}' = (\delta_2, -\delta_1)$. Nótese que este último vector es proporcional al (a_{11}, a_{12}).

DEMOSTRACION

Desde el punto del infinito de la dirección perpendicular a $\bar{\delta} = (\delta_1, \delta_2)$, esto es, de la dirección del vector $\bar{\delta}' = (\delta_2, -\delta_1)$, se pueden trazar dos tangentes a la parábola, una de las cuales es la recta del infinito. La otra tangente, que tiene la dirección de $\bar{\delta}'$, es por ello perpendicular a la dirección de la parábola y, entonces, su punto de tangencia es el vértice de la parábola. El eje es, según

la definición, la recta que une el vértice con el punto del infinito de la parábola; es decir, el eje es la recta que determinan los puntos de tangencia de las dos tangentes, a la parábola, desde el punto del infinito de la dirección de $\bar{\delta}'$, luego el eje es la polar (véase [309], 2) de dicho punto, es decir, la recta de ecuación $[0\ \delta_2\ -\delta_1]M\chi = O$.

Puede ser útil observar que, como, en det M, la suma de los productos de los elementos de su segunda fila por los respectivos adjuntos de la primera, $(\delta_0, \delta_1, \delta_2)$ con $\delta_0 = 0$, es cero, resulta que $a_{11}\delta + a_{12}\delta_2 = 0$, luego (a_{11}, a_{12}) es un vector ortogonal al $\bar{\delta} = (\delta_1, \delta_2)$, es decir, de la misma dirección que el $\bar{\delta}' = (\delta_2, -\delta_1)$.

La ecuación de la parábola respecto de su eje (eje de la x) y de la tangente en el vértice (eje de la y) no tiene término independiente (pues el origen es de la parábola), no tiene término en x^2 (pues el punto del infinito del eje x, punto $(0, 1, 0)$, pertenece a la parábola), no tiene término en xy (pues es parábola, o sea, su matriz de términos cuadráticos es singular) y no tiene término en y (pues la recta $x = 0$ es tangente a la parábola. Los otros dos términos (en y^2 y en x) son no nulos, pues de lo contrario la cónica sería degenerada). Por ello, la ecuación es del tipo $ay^2 + 2bx = 0$, para ciertos $a, b \in \mathbb{R}$ no nulos. Recurriendo a las invariantes métricas (véase [325]), se obtiene:

$$\det M = \det M' = -ab^2$$

$$a_{11} + a_{22} = a$$

luego la ecuación de la parábola en estos ejes es:

$$y^2 = -2\frac{b}{a}x = 2\sqrt{-\frac{\det M}{(a_{11} + a_{22})^3}}\,x$$

Finalmente, de esta ecuación se desprende, obviamente, que la parábola es simétrica respecto del «eje de la x», es decir, respecto de su eje.

EJERCICIO

Hallar la ecuación reducida de la siguiente cónica, determinando los ejes en los que la ecuación es la que se obtenga:

$$4x^2 - 4xy + y^2 + 2x + 4y - 10 = 0$$

RESOLUCION

Las matrices de la cónica y de sus términos cuadráticos son:

$$M = \begin{bmatrix} -10 & 1 & 2 \\ 1 & 4 & -2 \\ 2 & -2 & 1 \end{bmatrix} \qquad A = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}$$

Como $\det M = -25 \neq 0$ y $\det A = 0$, la cónica es una parábola y su ecuación reducida es (llamamos x_1 y x_2 a las correspondientes coordenadas):

$$x_2^2 = 2px_1 \quad \text{donde} \quad p = \sqrt{-\frac{\det M}{(a_{11}+a_{22})^3}} = \sqrt{\frac{25}{125}} = \frac{1}{\sqrt{5}}$$

Hallemos la referencia cartesiana pedida. El punto del infinito de la parábola se obtiene de (se llama t a la coordenada homogénea: la que otras veces denotamos por x_0):

$$\left.\begin{aligned} 4x^2 - 4xy + y^2 + 2xt + 4yt - 10t^2 = 0 \\ t = 0 \end{aligned}\right\} \quad \text{o sea} \quad \left.\begin{aligned} (2x - y)^2 = 0 \\ t = 0 \end{aligned}\right\}$$

luego dicho punto del infinito es el $(t, x, y) = (0, 1, 2)$. El punto del infinito de la dirección perpendicular es el $(0, 2, -1)$ y, por ello, el eje de la parábola (polar de dicho punto) es la recta

$$[0 \;\; 2 \;\; -1] M \begin{bmatrix} 1 \\ x \\ y \end{bmatrix} = O \quad \text{o sea} \quad y = 2x$$

Esta recta corta a la parábola en el punto $(x, y) = (1, 2)$, que es su vértice. La tangente en el vértice (perpendicular al eje) es $x + 2y = 5$. La referencia pedida tiene, pues:

- Origen en el vértice (1, 2).
- Primer vector unitario $\bar{e}_1 = (-1/\sqrt{5}, -2\sqrt{5})$ (*).
- Segundo vector unitario $\bar{e}_2 = \pm(2/\sqrt{5}, -1/\sqrt{5})$ (**).

β LAS CUADRICAS

En el espacio (tridimensional) afín, las superficies de segundo orden son los lugares de los puntos cuyas coordenadas (x_1, x_2, x_3), respecto de ejes cartesianos, satisfacen a una ecuación de segundo grado, que se podrá expresar de cualquiera de las formas siguientes:

$$\begin{aligned} a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 + \\ + 2b_1x_1 + 2b_2x_2 + 2b_3x_3 + c = 0 \end{aligned}$$

$$\sum_{i,j=1}^{3} a_{ij}x_ix_j + 2\sum_{h=1}^{3} b_hx_h + c = 0 \qquad (a_{21} = a_{12},\ a_{31} = a_{13},\ a_{32} = a_{23})$$

(*) Para dilucidar si el vector $\bar{e}_1$ es éste o su opuesto (ambos unitarios y de la dirección del eje, $y = 2x$), bastará con hallar algún punto de la parábola (como su intersección con $y = 2x + 1$) y ver a qué lado de la tangente en el vértice se encuentra dicho punto.

(**) Cualquiera de estos dos es válido.

$$[x_1\ x_2\ x_3]\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 2[b_1\ b_2\ b_3]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + [c] = O$$

$$[1\ x_1\ x_2\ x_3]\begin{bmatrix} c & b_1 & b_2 & b_3 \\ b_1 & a_{11} & a_{12} & a_{13} \\ b_2 & a_{21} & a_{22} & a_{23} \\ b_3 & a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = O$$

$$[1\ X^t]\begin{bmatrix} c & B^t \\ B & A \end{bmatrix}\begin{bmatrix} 1 \\ X \end{bmatrix} = O$$

El carácter de superficie de «segundo orden» se lo dan, obviamente, los sumandos de segundo orden o cuadráticas, esto es, los:

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3$$

Por ello, se exigirá que alguno de estos sumandos no sea cero, es decir, que la matriz $A = [a_{ij}]$ sea no nula.

14.4. LAS CUADRICAS; ESTUDIO PROYECTIVO

DEFINICION DE CUADRICA

[320]

En el espacio afín tridimensional, se llama *cuádrica* al lugar geométrico de los puntos cuyas coordenadas (x_1, x_2, x_3), respecto de una referencia cartesiana, verifican a una ecuación del tipo

$$\sum_{i,j=1}^{3} a_{ij}x_ix_j + 2\sum_{h=1}^{3} b_hx_h + c = 0 \qquad [1]$$

para ciertas a_{ij}, b_h, $c \in \mathbb{R}$ y donde la matriz $A = [a_{ij}]$ es simétrica y no nula. Con notación matricial, recurriendo a la matriz cuadrada $A = [a_{ij}]$, a la matriz columna $B = [b_h]$ y llamando X a la columna de las coordenadas x_1, x_2, x_3, la ecuación [1] se puede poner:

$$X^tAX + 2B^tX + [c] = O \quad \text{o} \quad [1\,X^t]M\begin{bmatrix} 1 \\ X \end{bmatrix} = O \quad \text{con} \quad M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \qquad [2]$$

A la matriz M, o a cualquiera de sus proporcionales no nulas, se la llama *matriz de la cuádrica* en los ejes dados; a la matriz A la llamaremos *matriz de los términos cuadráticos.*

NOTA: Un poco más adelante (véase [323]) comprobaremos que el carácter «cuadrático» de las cuádricas no depende de la referencia en la que se expresa su ecuación.

EJEMPLOS

Las cuádricas que ya conocemos, del capítulo anterior, satisfacen a la definición que acabamos de dar. Nos referimos a «las cinco cuádricas», cuyas ecuaciones reducidas son:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \qquad \frac{x^2}{a^2}\pm\frac{y^2}{b^2}-\frac{z^2}{c^2}=1, \qquad 2z=\frac{x^2}{p}\pm\frac{y^2}{q}$$

(elipsoide, hiperboloides y paraboloides, respectivamente; si el segundo miembro de la primera de ellas se sustituye por -1, se obtiene el «elipsoide imaginario»).

Hay también «cuádricas degeneradas», como las siguientes:

- $z^2=x^2+y^2$ (cono de revolución, alrededor del «eje z»).
- $x^2+y^2+z^2=0$ (punto $x=y=z=0$ o «cono imaginario» $z=\pm i\sqrt{x^2+y^2}$).
- $x^2+y^2=1$ (cilindro de revolución, alrededor del «eje z»).
- $y^2-2x=0$ (cilindro parabólico, de generatrices paralelas al «eje z»).

Hay cuádricas «más degeneradas» aún, como las:

- $x^2-y^2=0$ (pareja de planos: $x+y=0$ y $x-y=0$).
- $x^2+y^2=0$ (el «eje z» o par de planos imaginarios $y=\pm ix$).
- $x^2-1=0$ (pareja de planos: $x+1=0$ y $x-1=0$).
- $x^2+1=0$ (conjunto vacío o los «planos imaginarios» $x=\pm i$).
- $x^2=0$ (el plano $x=0$ doble, «contado dos veces»).

Según tendremos ocasión de comprobar más adelante, toda cuádrica es de alguno de los tipos que se acaban de considerar. Más exactamente, las cuádricas se clasifican en:

- Ordinarias (elipsoides, paraboloides e hiperboloides)
- Degeneradas { conos y cilindros; parejas de planos; planos dobles }

MATRICES DE UNA CUADRICA

En puridad, a la anterior definición [320], de cuádrica, se le puede poner el siguiente reparo: las cuádricas se han presentado como conjuntos de puntos cuando, en realidad vienen caracterizadas por sus ecuaciones y, en algunos casos, lo uno no se corresponde con lo otro. Así ocurre con las cuádricas $x^2+y^2+z^2+1=0$ y $x^2+4=0$: entendidas como conjuntos de puntos son, ambas, el vacío, luego se podría entender que eran iguales, aun cuando las ecuaciones que las definen son notablemente diferentes. Para salvar este escollo, se convendrá en que una cuádrica queda definida por su ecuación (en los ejes coordenados elegidos); más exactamente, una cuádrica se podrá caracterizar por una cualquiera de las matrices proporcionales a M (véase [320], [2]), lo que permite identificar a aquella con el conjunto de todas éstas matrices ρM, para $\rho\neq 0$.

[321]

En el espacio afín tridimensional y respecto de una referencia cartesiana, una cuádrica queda caracterizada por una matriz simétrica M, de tamaño 4×4, que tiene no nula la matriz que se obtiene al prescindir de sus primeras fila y columna (llamada submatriz de los términos cuadráticos). La cuádrica es equivalente al conjunto que forman todas las matrices (no nulas) proporcionales a M; de manera que:

$$\left(\text{Cuádrica de ecuación } [1\,X^t]M\begin{bmatrix}1\\X\end{bmatrix}=O^{(*)}\right)\simeq\{\lambda M/\lambda\in\mathbb{R},\ \lambda\neq0\}$$

Dos cuádricas son iguales si, y sólo si, sus matrices (en la misma referencia cartesiana) son proporcionales.

(*) Por X se denota a la columna de las coordenadas x_1, x_2, x_3 de un punto genérico del espacio.

EJERCICIO

Hallar una cuádrica tal que (respecto de unos ejes cartesianos $Oxyz$): 1) Su intersección con el plano $z=0$ es el par de bisectrices de los ejes Ox y Oy; 2) es simétrica respecto del plano $x=0$; 3) su intersección con el plano $y=0$ es la circunferencia con centro en (0, 0, 1) y radio unidad.

Compruébese que, para cada una de las cuádricas anteriores, hay dos planos paralelos al $x=0$ que la cortan según un par de rectas.

RESOLUCION

Llamando a_{ij}, b_h y c a los coeficientes de la ecuación de la cuádrica (como en [320]), para que sea simétrica respecto del plano $x=0$, ha de ser $a_{12}=a_{13}=b_1=0$. Para que su intersección con el plano $z=0$ sea $x^2-y^2=0$, $z=0$, ha de ser $b_2=c=0$ y $a_{22}=-a_{11}$. Para que su intersección con el plano $y=0$ sea $x^2+(z-1)^2=1$, $y=0$, ha de ser $a_{33}=a_{11}=-b_3$. Por tanto (dividiendo por a_{11} y llamando $p=a_{23}/a_{11}$), resulta que las cuádricas pedidas son las:

$$x^2-y^2+z^2+2pyz-2z=0\quad(\lambda\in\mathbb{R})$$

Al cortar esta cuádrica por el plano $x=h$ (h constante), se obtiene la cónica:

$$-y^2+z^2+2pyz-2z+h^2=0\quad\text{(además de } x=h)$$

Esta cónica será degenerada (un par de rectas) si su determinante es nulo, esto es, si $1-h^2(1+p^2)=0$. Por tanto, los planos

$$x=h=\pm\frac{1}{\sqrt{1+p^2}}$$

cortan a la cuádrica según un par de rectas.

CAMBIO DE COQRDENADAS

[322]

En el espacio afín tridimensional, se considera la cuádrica que, en una cierta referencia cartesiana (coordenadas x_1, x_2, x_3), tiene a M por matriz, es decir, que admite por ecuación a:

$$[1\,X^t]M\begin{bmatrix}1\\X\end{bmatrix}=O \quad \text{siendo} \quad X=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} \quad \text{y} \quad M=\begin{bmatrix}c & B^t\\B & A\end{bmatrix} \qquad [1]$$

(donde: $A \neq 0$ simétrica 3×3; B columna 3×1; c término independiente).

Supongamos, también, que se adopta una nueva referencia cartesiana (coordenadas x'_1, x'_2, x'_3) y que el correspondiente cambio de coordenadas es (véase [255]):

$$\begin{bmatrix}1\\X\end{bmatrix}=Q^*\begin{bmatrix}1\\X\end{bmatrix} \quad \text{con} \quad Q^*=\begin{bmatrix}1 & O\\C & Q\end{bmatrix} \qquad [2]$$

(C es la columna 3×1 de coordenadas del nuevo origen; Q es la matriz 3×3 correspondiente al cambio de base).

Entonces, en la nueva referencia, la cuádrica tiene por matriz a la M' siguiente, que es congruente con M (véase [114]):

$$M'=Q^{*t}MQ^* \quad \text{o} \quad M'=\begin{bmatrix}c' & B'^t\\B' & A'\end{bmatrix}=\begin{bmatrix}1 & C^t\\O & Q^t\end{bmatrix}\begin{bmatrix}c & B^t\\B & A\end{bmatrix}\begin{bmatrix}1 & O\\C & Q\end{bmatrix} \qquad [3]$$

Y, en particular, las matrices de los términos cuadráticos, A y A', también son congruentes, pues están relacionadas entre sí mediante:

$$A'=Q^tAQ$$

COMPROBACION

Vale aquí los dicho en la demostración de [303] (donde allí dice cónica, dígase aquí cuádrica).

EJERCICIO

En el espacio afín euclídeo y respecto de ejes rectangulares $Oxyz$, se considera la cuádrica:

$$2x^2+2y^2+2z^2-2xy-2xz+x-y-z+1=0$$

Se cambian los anteriores ejes por otros que tienen origen en el punto (0, 1, 0) y cuyos vectores son los unitarios de las bisectrices de los ángulos yOz, zOx, xOy. Hallar la ecuación de la cuádrica en los nuevos ejes.

RESOLUCION

El cambio de coordenadas (llamamos x', y', z' a las nuevas) es:

$$x = (y' + z')/\sqrt{2}$$
$$y = (z' + x')/\sqrt{2} + 1$$
$$z = (x' + y')/\sqrt{2}$$

Llevando estas expresiones a la ecuación de la cuádrica, se obtiene:

$$2x'^2 + y'^2 + z'^2 + \sqrt{2}y' + 3 = 0$$

INVARIANTES PROYECTIVO Y AFIN

[323]

En el espacio tridimensional afín, el hecho de que una superficie sea cuádrica, esto es, con ecuación de segundo grado, es independiente de la referencia cartesiana en la que aquella se expresa. Aún más:

Supóngase que, en una cierta referencia cartesiana del espacio, M es la matriz simétrica de una cuádrica y que A es su submatriz (simétrica) de términos cuadráticos. Supóngase que, en otra referencia, las referidas matrices son M' y A'. Entonces se verifica que las signaturas de $\pm M$ y $\pm A$ son invariantes; es decir:

$$\text{Sig}\, M' = \text{sig}\,(\pm M)\ (*), \quad \text{en particular} \quad [\det M' = 0] \Leftrightarrow [\det M = 0]$$
$$\text{Sig}\, A' = \text{sig}\,(\pm A), \quad \text{en particular} \quad [\det A' = 0] \Leftrightarrow [\det A = 0]$$

(*) Como consecuencia, se verifica que det M y det M' tienen el mismo signo (cosa que puede no ocurrir con det A).

DEMOSTRACION

Vale aquí lo dicho en la demostración de [304] (donde allí dice plano, dígase aquí espacio tridimensional).

La observación del final es consecuencia inmediata de que $\det M$ es positivo si, y sólo si, $\text{sig}\, M$ vale (4, 0), (2, 2) ó (0, 4).

EJERCICIO

De una cuádrica se sabe que, en cierta referencia cartesiana del espacio, su matriz de términos cuadráticos, A, tiene signatura $\text{sig}\, A = (3, 0)$.

También se sabe que, en otra referencia, la cuádrica admite ecuación del tipo

$$x^2 + \widetilde{ay^2} + bz^2 + c = 0$$

Si la cuádrica tiene más de un punto, hállense los signos de a, b y c.

RESOLUCION

Como la signatura de la matriz de términos cuadráticos no varía al cambiar de referencia, resulta que

$$\text{sig}\, A = (3, 0) \quad \Rightarrow \quad a \text{ y } b \text{ son positivos}$$

Si c también fuese positivo, la ecuación $x^2 + ay^2 + bz^2 + c = 0$ no se verificaría para ningún (x, y, z), que no es cierto, pues la cuádrica tiene puntos. Si c fuese nulo, la ecuación $x^2 + ay^2 + bz^2 + c = 0$ sólo se verificaría para $(x, y, z) = (0, 0, 0)$, lo que tampoco es cierto, pues la cuádrica tiene más de un punto. Por tanto, c ha de ser negativo.

ECUACION HOMOGENEA DE UNA CUADRICA

El recurso a las coordenadas homogéneas nos va a resultar muy provechoso, sobre todo en aquellas cuestiones en las que intervengan los puntos del infinito de las cuádricas.

Según dijimos en [237], dada una referencia cartesiana en el espacio afín, si (x_1, x_2, x_3) es el sistema de coordenadas (absolutas) de un punto, entonces cualquiera de las cuaternas

$$(X_0, X_1, X_2, X_3) = (\rho, \rho x_1, \rho x_2, \rho x_3) \quad \text{para} \quad \rho \neq 0$$

es un sistema de coordenadas homogéneas del punto en cuestión. En lo que sigue, denotaremos por X y $\mathcal{X}$ a las columnas de coordenadas absolutas y homogéneas de un punto genérico del espacio; así, pues, para algún $\rho \neq 0$ es:

$$\mathcal{X} = \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \rho \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \rho \begin{bmatrix} 1 \\ X \end{bmatrix}$$

Es decir, $x_i = X_i/X_0$ (para $i = 1, 2, 3$), siendo $X_0 \neq 0$ arbitrario.

Recuérdese que los puntos del infinito se introdujeron como unos nuevos elementos, que también tenían coordenadas homogéneas, siendo éstas las cuaternas $(0, X_1, X_2, X_3)$; es decir, los puntos impropios o del infinito quedaron caracterizados por la relación $X_0 = 0$. Con ello, la única cuaterna (X_0, X_1, X_2, X_3) que no es sistema de coordenadas de un punto (propio o impropio) es la cuaterna nula ($X_0 = X_1 = X_2 = X_3 = 0$ no representa a ningún punto).

[324]

En el espacio afín tridimensional y respecto de una referencia cartesiana (coordenadas absolutas x_1, x_2, x_3; coordenadas homogéneas X_0, X_1, X_2, X_3), la cuádrica cuya ecuación absoluta es [1] tiene [2] como ecuación homogénea, siendo:

$$\sum_{i,j=1}^{3} a_{ij}x_ix_j + 2\sum_{h=1}^{3} b_hx_h + c = 0 \qquad [1]$$

$$\sum_{i=1}^{3} a_{ij}X_iX_j + 2\sum_{h=1}^{3} b_hX_0X_h + cX_0^2 = 0 \qquad [2]$$

Esta ecuación [2], al expresarla matricialmente, queda:

$$[X_0X_1X_2X_3]\begin{bmatrix} c & b_1 & b_2 & b_3 \\ b_1 & a_{11} & a_{12} & a_{13} \\ b_2 & a_{21} & a_{22} & a_{23} \\ b_3 & a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = O \quad \text{o} \quad X^tMX = O$$

NOTA: Respecto de la notación, se suele poner $c = a_{00}$ y $b_h = a_{0h} = a_{h0}$, con lo que $M = [a_{ij}]$ (para $i, j = 0, 1, 2, 3$) y la ecuación [2] queda $\Sigma a_{ij}X_iX_j = 0$ (la suma para $i, j = 0, 1, 2, 3$).

OBSERVACION

Hay que tener en cuenta que, en la anterior ecuación $X^tMX = O$ de una cuádrica, la matriz M no puede ser cualquiera ya que, según se dijo anteriormente (véase [320]), se exige que los elementos a_{ij} (para $i, j = 1, 2, 3$) no sean todos cero, esto es, que la submatriza $A = [a_{ij}]$, de los términos cuadráticos, no sea nula. Nótese que si A fuese nula, las ecuaciones [1] y [2] serían:

$$2\sum_{h=1}^{3} b_hx_h + c = 0 \quad \text{y} \quad \left[2\sum_{h=1}^{3} b_hX_h + cX_0\right]X_0 = 0$$

En el espacio afín (sin puntos del infinito) estas ecuaciones lo son de un plano. En el espacio ampliado y utilizando coordenadas homogéneas, la última ecuación representa, no sólo a los puntos del plano citado, sino también a los puntos del plano del infinito ($X_0 = 0$). Resulta, entonces, que al obligar a que sea $A \neq O$ se esta rechazando, de entre las cuádricas, al lugar geométrico que forman los puntos de un plano a los que se unirían, si se considera el espacio ampliado, los puntos del infinito.

EJERCICIO

En el espacio afín y respecto de una referencia cartesiana (coordenadas x, y, z), se consideran las cuádricas de ecuaciones:

$$2x^2 + 3y^2 + z^2 - 3xy + 2yz + z + 7 = 0 \qquad [1]$$

$$2x^2 + 3y^2 - z^2 + 5xy - 2xz + 3y - 2z + 1 = 0 \qquad [2]$$

1. Compruébese que la cuádrica [1] carece de puntos en el infinito.
2. Hállense todas las cuádricas que tienen los mismos puntos en el infinito que la cuádrica [2].

RESOLUCION

Utilicemos coordenadas homogéneas, que llamaremos (t, x, y, z).

1. Los puntos del infinito de la cuádrica [1] son las soluciones del sistema

$$\left.\begin{aligned} 2x^2 + 3y^2 + z^2 = 0 \\ t = 0 \end{aligned}\right\}$$

   Como su única solución es $(t, x, y, z) = (0, 0, 0, 0)$ y esta cuaterna no es sistema de coordenadas de ningún punto, la cuádrica no tiene puntos en el infinito.
2. Los puntos del infinito de dos cuádricas (la [2] y otra) son los mismos si ambas tienen los mismos términos cuadráticos (lo que queda de la ecuación al hacer $t = 0$), las cuádricas que tienen los mismos puntos en el infinito que la [2] son, pues, las:

$$2x^2 + 3y^2 - z^2 + ax + by + cz + d = 0 \qquad (a, b, c, d \in \mathbb{R})$$

EJERCICIO

En el espacio afín y respecto de una referencia cartesiana (coordenadas absolutas x_1, x_2, x_3; coordenadas homogéneas X_0, X_1, X_2, X_3) se considera la cuádrica C cuya ecuación es

$$\sum_{i,j=1}^{3} a_{ij} X_i X_j + 2 \sum_{h=1}^{3} b_h X_0 X_h + c = 0 \qquad [1]$$

1. Hallar el lugar geométrico de los puntos del infinito de C (que se llama cónica del infinito de C).
2. Hallar una condición necesaria y suficiente para que C no tenga puntos en el infinito.
3. Si C tiene puntos en el infinito, hallar la ecuación del cono que proyecta a dichos puntos del infinito desde el origen de coordenadas.

RESOLUCION

1. La intersección de [1] con $X_0 = 0$ conduce a $\Sigma a_{ij} X_i X_j = 0$; más exactamente:

$$\begin{cases} a_{11}X_1^2 + a_{22}X_2^2 + a_{33}X_3^2 + 2a_{12}X_1X_2 + 2a_{13}X_1X_3 + 2a_{23}X_2X_3 = 0 \\ X_0 = 0 \end{cases} \qquad [2]$$

2. La primera de las ecuaciones de [2] debe carecer de solución (salvo la $X_1 = X_2 = X_3 = 0$), lo que equivale a decir que la matriz $A = [a_{ij}]$, de tamaño 3×3, sea definida, esto es $\text{sig}\, A = (3, 0)$ ó $(0, 3)$, lo que se expresará poniendo $\text{sig}\, A = \{3, 0\}$.
3. Al proyectar [2] desde el origen de coordenadas se obtiene el cono cuya ecuación, es obviamente, la primera de las ecuaciones [2].

INTERSECCION DE CUADRICA Y RECTA

[325]

En el espacio afín tridimensional ampliado (con los puntos del infinito), una cuádrica y una recta se cortan en dos, uno o ningún punto y se dice que la recta es, respectivamente, *secante*, *tangente* o *exterior* a la cuádrica; también cabe la posibilidad de que la recta esté incluida en la cuádrica (se dirá entonces que la cuádrica es reglada).

Supongamos que, en una cierta referencia cartesiana (coordenadas homogéneas X_0, X_1, X_2, X_3), la cuádrica y la recta admiten por ecuaciones homogéneas a las:

$$X^t M X = O, \text{ donde } \begin{cases} X = [X_i] \text{ es la columna de coordenadas } (4 \times 1) \\ M = \text{matriz (simétrica) de la cuádrica} \end{cases}$$

$$X = \lambda\alpha + \mu\beta, \text{ donde } \begin{cases} X,\ \alpha \text{ y } \beta \text{ son columnas de} \\ \text{coordenadas del punto genérico} \\ \text{y de dos puntos fijos de la recta} \end{cases}$$

Entonces los puntos de su intersección se obtienen para los (λ, μ), parámetros, soluciones de, que son solución de la ecuación:

$$(\alpha^t M \alpha)\lambda^2 + 2(\alpha^t M \beta)\lambda\mu + (\beta^t M \beta)\mu^2 = O \qquad [1]$$

La recta es tangente a la cuádrica si se verifica la relación:

$$(\alpha^t M \beta)^2 = (\alpha^t M \alpha)(\beta^t M \beta) \qquad [2]$$

La recta estará incluida en la cuádrica si los tres coeficientes de la ecuación [1] son nulos, esto es, si $\alpha^t M \beta = \alpha^t M \alpha = \beta^t M \beta = O$.

COMPROBACION

Vale aquí lo dicho en la comprobación de [306] (donde allí dice cónica, dígase aquí cuádrica).

OBSERVACION

Según lo que se acaba de decir, en el espacio afín ampliado se verifica que una recta y una cuádrica tienen dos puntos en común, que pueden ser reales y

distintos, pueden coincidir o pueden ser imaginarios. Si en lugar de movernos en el espacio ampliado, sólo consideramos los puntos propios (pero no los del infinito), lo anterior puede ser falso ya que algunos de los puntos anteriores pueden ser de los del infinito y, por ello, ahora se nos escaparían.

Así ocurre, por ejemplo, al buscar la intersección de la recta y la cuádrica siguientes:

$$\begin{cases} x = 2 \\ y = 1 \end{cases} \qquad 3x^2 - 8y^2 + 2z = 0$$

Al resolver este sistema nos encontramos con la única solución $(x, y, z) = (2, 1, -2)$. Si consideramos también los puntos del infinito y manejando coordenadas homogéneas (t, x, y, z), el anterior sistema se convierte en el

$$\begin{cases} x = 2t \\ y = t \end{cases} \qquad 3x^2 - 8y^2 + 2zt = 0$$

cuyas soluciones son $(t, x, y, z) = \rho(1, 2, 1, -2)$, que es la hallada antes, y $(t, x, y, z) = \rho(0, 0, 0, 1)$, que es un punto del infinito.

INTERSECCION DE CUADRICA Y PLANO

[326]

En el espacio afín tridimensional ampliado, la intersección de una cuádrica y un plano es una cónica.
Según que esta cónica sea ordinaria y real, sea ordinaria e imaginaria o sea degenerada, se dice que el plano es, respectivamente, *secante*, *exterior* o *tangente* a la cuádrica. En el caso de tangencia, se dice que el punto singular de la intersección es el punto de tangencia del plano.

COMPROBACION

El plano dado puede tomarse como plano $x_3 = 0$ de una referencia cartesiana $(O, \bar{e}_1, \bar{e}_2, \bar{e}_3)$ (coordenadas cartesianas x_1, x_2, x_3); la ecuación de la cuádrica en esta referencia será (para ciertos coeficientes a_{ij}, b_h y c):

$$\sum_{i,j=1}^{3} a_{ij} x_i x_j + \sum_{h=1}^{3} b_h x_h + c = 0 \qquad [1]$$

Por tanto, la intersección de esta cuádrica con el plano $(x_3 = 0)$ es la curva de este plano que, en la referencia $(O; \bar{e}_1, \bar{e}_2)$ admite por ecuación a

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c = 0$$

que es una cónica, como había que comprobar.

EJERCICIO

Sean dados el plano y la cuádrica siguientes:

$$\pi : x - 2y + 2z = 2$$
$$C : x^2 - 2y^2 + 3z^2 + 4xy + 2yz + 2x - 2z + 1 = 0$$

obtener la ecuación de su cónica de intersección en una referencia cartesiana de su plano, que se podrá elegir.

RESOLUCION

En el plano π tomemos el punto $P(0, 0, 1)$ y sean $\vec{u} = (2, 0, -1)$ y $\vec{v} = (0, 1, 1)$ dos vectores de la dirección del plano. Si recurrimos también al vector $\vec{w} = (0, 0, 1)$, como $(\vec{u}, \vec{v}, \vec{w})$ es base de $\mathbb{R}^3$, podemos recurrir a la referencia cartesiana $(A; \vec{u}, \vec{v}, \vec{w})$; a las coordenadas en esta referencia las llamaremos (x', y', z'). El correspondiente cambio de coordenadas es:

$$\begin{bmatrix} 1 \\ x \\ y \\ z \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{array}\right] \begin{bmatrix} 1 \\ x' \\ y' \\ z' \end{bmatrix}$$

abreviadamente

$$\begin{bmatrix} 1 \\ X \end{bmatrix} = Q^* \begin{bmatrix} 1 \\ X' \end{bmatrix}$$

Llamando M a la matriz de la cuádrica en la referencia dada, la matriz M' en la nueva referencia será $M' = Q^{*t} M Q^*$. Como M es la matriz que figura a continuación, al realizar el producto anterior, se obtiene la siguiente M':

$$M = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 1 & 2 & 0 \\ 0 & 2 & -2 & 1 \\ 1 & 0 & 1 & 3 \end{bmatrix} \quad ; \quad M' = \begin{bmatrix} 2 & 0 & 3 & 2 \\ 0 & 7 & 0 & -3 \\ 3 & 0 & 3 & 4 \\ 2 & -3 & 4 & 3 \end{bmatrix}$$

La ecuación de la cónica $C \cap \pi$ en la referencia $(A; \vec{u}, \vec{v})$ de su plano se obtendrá de hacer $z' = 0$ en la ecuación de la cuádrica C en la referencia $(A; \vec{u}, \vec{v}, \vec{w})$. Así, pues, la ecuación pedida es:

$$7x'^2 + 3y'^2 + 6y' + 2 = 0$$

CUADRICAS ORDINARIAS Y DEGENERADAS

[327]

I. En el espacio afín tridimensional ampliado y respecto de una referencia cartesiana, se considera la cuádrica cuya matriz es M (simétrica 4×4). Se dice que la cuádrica es *ordinaria* o que es *degenerada* según que M sea regular (det $M=0$) o singular (det $M\neq0$) [1], respectivamente.

II. Si la cuádrica es degenerada, entonces se llaman *puntos singulares* de la misma a aquellos puntos cuya matriz X (columna de sus coordenadas homogéneas X_0, X_1, X_2, X_3) es una solución no nula de la ecuación $MX=O$.
Si P_s es un punto singular de una cuádrica degenerada, entonces ésta es la reunión de una familia de rectas que pasan por P_s.

III. Cuando la cuádrica es degenerada (rang $M<4$), se verifica que: 1) si rang $M=3$, entonces tiene un solo punto singular P_s y la cuádrica es un *cono* [2] o un *cilindro* con vértica en P_s; 2) si rang $M=2$, entonces hay una recta de puntos singulares y la cuádrica es un par de planos que se cortan según esa recta; y 3) si rang $M=1$, entonces hay un plano de puntos singulares y la cuádrica es ese plano doble.

([1]) Esta definición no depende de la referencia cartesiana, ya que la anulación de det M es invariante frente a los cambios de referencia (véase [323]).

([2]) Más exactamente, un cono cuadrático: lugar geométrico que engendran las rectas que pasan por un punto fijo (vértice) y por un puto variable que recorre una cónica ordinaria no coplanaria con el vértice. Si el vértice está en el infinito, en lugar de cono, se llama cilindro.

DEMOSTRACION

Para comprobar el apartado II vale lo dicho en la demostración de [307] (donde allí dice cónica, dígase aquí cuádrica). Hablaremos, pues, sólo del apartado III.

Como el sistema $MX=O$ es homogéneo, sus soluciones constituyen un espacio vectorial: si las columnas $X=\alpha_1, \ldots, X=\alpha_p$ forman una base de este espacio (donde $p=4-\text{rang}\,M$), entonces todas las soluciones son las $X=\lambda_1\alpha_1+\cdots+\lambda_p\alpha_p$ (para $\lambda_1, \ldots, \lambda_p\in\mathbb{R}$). Por tanto, si rang $M=3$, las soluciones son las $X=\lambda_1\alpha_1$ (para $\lambda_1\in\mathbb{R}$), que representa a un solo punto singular (el que tiene por columna de coordenadas homogéneas a α_1); si rang $M=2$, hay una recta de puntos singulares, ya que las soluciones son las $X=\lambda_1\alpha_1+\lambda_2\alpha_2$ (para $\lambda_1, \lambda_2\in\mathbb{R}$), que representa a la recta que pasa por los puntos de coordenadas α_1 y α_2; si rang $M=1$, hay un plano de puntos singulares, ya que las soluciones son ahora las $X=\lambda_1\alpha_1+\lambda_2\alpha_2+\lambda_3\alpha_3$ (para $\lambda_1, \lambda_2, \lambda_3\in\mathbb{R}$).

En el caso de rang $M=3$, veamos que la cuádrica es un cono o un cilindro, que tiene por vértice a su único punto singular P_s. Para ello, traslademos los

ejes coordenados de modo que P_s pasa a ser el nuevo origen (si P_s fuese un punto del infinito, en lugar de trasladar los ejes, se tomaría uno de ellos con la dirección de P_s; en este caso se podría razonar análogamente a como aquí se hace). Como ahora la solución de la ecuación $M\chi = O$ es la columna de las coordenadas de P_s (1, 0, 0, 0), resulta que la matriz M ha de tener nula su primera columna (como es simétrica, también tiene nula su primera fila) y la ecuación de la cuádrica sería (se llaman x_1, x_2, x_3 a las coordenadas absolutas):

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 = 0$$

Esta cuádrica está formada por rectas que pasan por el origen ya que, si un punto (x_1, x_2, x_3) verifica a su ecuación, también la verifican todos los puntos $(\lambda x_1, \lambda x_2, \lambda x_3)$, para $\lambda \in \mathbb{R}$, que son los alineados con aquél y el origen. Además esta cuádrica es un «cono cuadrático», esto es, la superficie que resulta de proyectar desde el origen una cónica ordinaria situada en un plano que no pasa por el origen. En efecto al cortar dicha cuádrica por el plano $x_3 = 1$ (por ejemplo), nos encontramos con la cónica que, en la referencia obvia(*), tiene por ecuación a:

$$a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2 + 2a_{13}x_1 + 2a_{23}x_2 + a_{33} = 0$$

Nótese que esta cónica es, realmente, no degenerada pues su matriz, que es $A = [a_{ij}]$ (para $i, j = 1, 2, 3$), es regular ya que la matriz de la cuádrica es

$$M = \begin{bmatrix} 0 & O \\ O & A \end{bmatrix} \quad \text{y} \quad 3 = \operatorname{rang} M = \operatorname{rang} A$$

Para los casos en los que $\operatorname{rang} M$ vale 2 ó 1, se puede proceder de modo análogo: si $\operatorname{rang} M = 2$, se toma a la recta de puntos singulares como nueva recta $x_2 = x_3 = 0$ (eje de la x_1); si $\operatorname{rang} M = 1$, se toma al plano de los puntos singulares como nuevo plano $x_1 = 0$. En estos casos, las cosas resultan más simples que en el anterior, el de $\operatorname{rang} M = 3$.

EJERCICIO

Pruébese que si una cuádrica incluye a tres rectas (distintas) que concurren en un punto, entonces la cuádrica es degenerada.

RESOLUCION

Si las tres rectas no son coplanarias, tomándolas como ejes de una referencia cartesiana, la ecuación de la cuádrica, respecto de dichos ejes, será del tipo

$$axy + bxz + cyz = 0$$

(*) Si la referencia que se ha utilizado en el espacio es la $(P_s; \vec{u}_1, \vec{u}_2, \vec{u}_3)$, la referencia a la que aquí nos referimos es la $(P, \vec{u}_1, \vec{u}_2)$, donde $P(0, 0, 1)$.

ya que debe verificarse para todos los puntos $(x, 0, 0)$, $(0, y, 0)$ y $(0, 0, z)$. Como la matriz de esta cuádrica es la:

$$M = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & a & 0 & c \\ 0 & b & c & 0 \end{bmatrix} \quad \text{y} \quad \det M = 0$$

se trata de una cuádrica degenerada.

Si las tres rectas son coplanarias, tomemos dos de ellas como ejes de una referencia cartesiana (eje de la x y eje de la y), con lo que la tercera recta tendrá ecuaciones del tipo $y = mx$, $z = 0$ (para cierto $m \neq 0$). En tal referencia, la cuádrica tendrá ecuación del tipo

$$z(ax + by + cz + d) = 0$$

ya que debe verificarse para todos los puntos $(x, 0, 0)$, $(0, y, 0)$ y $(x, mx, 0)$. La cuádrica está, pues, formada por dos planos:

$$z = 0 \quad \text{(que contiene a las dos rectas)}$$
$$ax + by + cz + d = 0$$

y es entonces degenerada (véase que su matriz tiene rango 2, a lo sumo).

EJERCICIO

Considérese la cuádrica que, respecto de una cierta referencia cartesiana, tiene por matriz a

$$M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \quad \text{donde} \quad \begin{cases} A \text{ matriz simétrica } 3 \times 3 \\ B \text{ matriz columna } 3 \times 1 \\ c \text{ término independiente} \end{cases}$$

Hállese, cuando exista, un valor $c' \in \mathbb{R}$ que, al sustituirlo por c, haga que la cuádrica sea degenerada.

RESOLUCION

Sea M' la matriz que resulta de sustituir c' por c en M. Ha de ser $\det M' = 0$. Recurriendo a las propiedades de la determinante y llamando M_0 a la matriz que se obtiene al poner 0 en lugar de c en M, será:

$$0 = \det M' = \begin{vmatrix} c' & B^t \\ B & A \end{vmatrix} = \begin{vmatrix} 0 & B^t \\ B & A \end{vmatrix} + \begin{vmatrix} c' & B^t \\ O & A \end{vmatrix} = \det M_0 + c' \ (\det A)$$

Luego el problema tiene solución si $\det A \neq 0$ y ésta es

$$c' = -(\det M_0) : (\det A)$$

CUADRICAS REALES E IMAGINARIAS

[328]

Una cuádrica que tenga algún punto (uno al menos) se dice que es real; si carece de puntos, se la llama imaginaria.

En el espacio afín ampliado y respecto de una referencia cartesiana, la cuádrica cuya matriz es M (simétrica, 4×4) es imaginaria o es real según que, respectivamente, M sea definida o no lo sea, esto es, según que la signatura de M sea $\{4, 0\}$ (*) o no lo sea.

NOTA: Cuando decimos que la cuádrica carece de puntos (M es definida), nos referimos, tanto, a puntos propios como a puntos del infinito.

(*) Para señalar que sig$M$ vale (3, 0) ó (0, 3), se pondrá sig$M$ = {3, 0}.

DEMOSTRACION

Vale aquí lo dicho en la demostración de [308] (donde allí se dice plano, dígase aquí espacio tridimensional).

EJERCICIO

En el espacio afín y respecto de una referencia cartesiana se considera la cuádrica C cuya ecuación es:

$$X^tAX + 2B^tX + [c] = O \qquad [1]$$

(A matriz simétrica 3×3; B matriz columna 3×1: X es la columna (3×1) de coordenadas; $c\in\mathbb{R}$). Recurriendo a la intersección de C con una recta variable que pase por el origen, compruébese que:

$$[C \text{ es imaginaria}] \Leftrightarrow [BB^t - cA \text{ es definida negativa}]$$

RESOLUCION

Sea $\bar{u}(u_1, u_2, u_3)\in\mathbb{R}^3$ un vector cualquiera. La recta que pasa por el origen y tiene la dirección de $\bar{u}$ está formada por los puntos cuya columna de coordenadas es $X = \lambda U$, para $\lambda\in\mathbb{R}$ y donde U es la columna de coordenadas de $\bar{u}$. Poniendo $X = \lambda U$ en [1], los puntos de intersección de esta recta con C resultan ser los que se obtienen al tomar para λ las raíces de

$$\lambda^2(U^tAU) + 2\lambda(B^tU) + [c] = O \qquad [2]$$

La cuádrica C será imaginaria si la ecuación [2] (en la incógnita λ) no tiene raíces reales, esto es, si y sólo si su discriminante es negativo:

$$(B^tU)^2 - (U^tAU)c < 0 \quad \text{para toda } U\in\mathcal{M}_{3\times1}$$

Esta relación se puede poner en la forma

$$U^t[BB^t - cA]U < 0 \quad \text{para toda } U\in\mathcal{M}_{3\times1}$$

esto es, si la matriz $BB^t - cA$ es definida negativa.

TANGENTES A UNA CUADRICA

[329]

En el espacio afín tridimensional y respecto de una referencia cartesiana (sea X la columna de las coordenadas homogéneas de un punto genérico), se considera la cuádrica cuya matriz es M (esto es, cuya ecuación homogénea es $X^tMX = O$); se considera también el punto A cuya columna de coordenadas homogéneas es α.

I. Se verifica que las rectas tangentes a la cuádrica que pasan por A engendran lo que se llama el *cono circunscrito* a la cuádrica desde A, que es una cuádrica degenerada y tiene por ecuación a:

$$(\alpha^tMX)^2 = (\alpha^tM\alpha)(X^tMX) \qquad [1]$$

II. Si A es un punto ordinario de la cuádrica[1], entonces el «cono circunscrito» desde A es un plano (doble), que es el *plano tangente* a la cuádrica en el punto A y tiene por ecuación homogénea a la:

$$\alpha^tMX = O \qquad [2]$$

(en este caso, a este plano se le llama, también, plano polar de A).

III. Si A no pertenece a la cuádrica, el «cono circunscrito» desde A es un cono o un par de planos, según que la cuádrica sea, respectivamente, ordinaria o cono [2]. Los puntos de contacto de las tangentes desde A (que no es de la cuádrica) forman una cónica, que es la intersección de la cuádrica con un plano, que se llama *plano polar* del punto A respecto de la cuádrica, y tiene por ecuación homogénea a $\alpha^tMX = O$.

---

(1) Si A fuese un punto singular de la cuádrica, la ecuación [1] se verificaría para todo punto del espacio.

(2) Si la cuádrica fuese un par de planos o un plano doble el «cono circunscrito» sería un plano doble o todo el espacio, respectivamente. Estos casos «patológicos» tienen poco interés.

DEMOSTRACION

I. Según ya sabemos (véase [325], [2]), una recta AB es tangente a la cuádrica si, llamando β a la columna de coordenadas homogéneas de B, se verifica que:

$$(\alpha^tM\beta)^2 = (\alpha^tM\alpha)(\beta^tM\beta)$$

Si aquí se sustituye la columna fija β por la columna variable X de las coordenadas homogéneas de un punto genérico del espacio, se obtiene, pues, la ecuación del conjunto de las rectas tangentes a la cuádrica desde A, que es entonces

$$(\alpha^tMX)^2 = (\alpha^tM\alpha)(X^tMX) \qquad [1]$$

II. Si A es un punto de la cuádrica, entonces es $\alpha^t M\alpha = O$ y, por ello, la ecuación [1] queda reducida a $(\alpha^t M\chi)^2 = O$, que es el lugar de ecuación $\alpha^t M\chi = O$ contado dos veces; este lugar es un plano (aquel cuyos coeficientes forman la fila $\alpha^t M$, que no es nula, ya que A no es un punto singular de la cuádrica).

Para concluir hay que comprobar que este plano, que contiene a todas las rectas tangentes a la cuádrica en A, es plano tangente según la definición dada anteriormente, en [326], esto es, su intersección con la cuádrica debe ser un par de rectas que pasan por A. Para ello, cambiemos los ejes coordenados (coordenadas absolutas x_1, x_2, x_3) de manera que A sea el nuevo origen y el plano [2] sea el $x_3 = 0$: en esta referencia, es evidente que la matriz de la cuádrica es del tipo

$$M = \begin{bmatrix} 0 & 0 & 0 & b_3 \\ 0 & a_{11} & a_{12} & a_{13} \\ 0 & a_{21} & a_{22} & a_{23} \\ b_3 & a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Por tanto, la intersección de la cuádrica con el plano $x_3 = 0$ es la cónica

$$\left.\begin{aligned} a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 0 \\ x_3 = 0 \end{aligned}\right\}$$

que es, evidentemente, un par de rectas que pasan por A.

III. Si comprobamos que, en una determinada referencia cartesiana, el rango de la matriz del cono circunscrito [1] es una unidad menor que el rango de la matriz de la cónica dada, resultará que si ésta es ordinaria aquél será cono y que si ésta es cono aquél será un par de planos. En el próximo ejercicio se puede encontrar la anterior relación entre los rangos.

Los puntos de contacto de las tangentes son la intersección del cono circunscrito, ecuación [2], con la cuádrica, de ecuación $\chi^t M\chi = O$. Como, para los puntos de tangencia, se verifican estas dos ecuaciones, también se verifican las que se deducen de ellas, como la $\alpha^t M\chi = O$, que representa a un plano (polar). Por tanto, los puntos de contacto de las tangentes desde A a la cuádrica son los de intersección de ella con el plano polar $\alpha^t M\chi = O$.

EJERCICIO

Sean dadas una cuádrica y un punto P que no está en ella; tómense unos ejes cartesianos con origen en P y supóngase que, en ellos, la matriz de la cuádrica es

$$M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \quad \text{donde} \quad \begin{cases} A = [a_{ij}] \text{ es matriz simétrica } 3\times 3 \\ B = [b_h] \text{ es matriz columna } 3\times 1 \\ c \in \mathbb{R} \text{ es el término independiente} \end{cases}$$

Hallar la matriz M_c del cono circunscrito a la cuádrica desde P, en la anterior referencia. Comprobar que

$$\operatorname{rang} M_c = (\operatorname{rang} M) - 1$$

RESOLUCION

Llamando x_1, x_2, x_3 a las coordenadas en estos ejes y según se acaba de obtener en [329], [1], la ecuación del cono circunscrito desde el punto $P(0, 0, 0)$ es:

$$(c + b_1x_1 + b_2x_2 + b_3x_3)^2 = c\left(\sum_{i,j=}^{3} a_{ij}x_ix_j + 2\sum_{h=1}^{3} b_hx_h + c\right)$$

Simplificando aquí, se obtiene la ecuación buscada o, lo que es equivalente, la matriz M_c:

$$\sum_{i,j=1}^{3} (b_ib_j - ca_{ij})x_ix_j = 0 \qquad M_c = \left[\begin{array}{c|c} 0 & O \\ \hline O & BB^t - cA \end{array}\right]$$

Obsérvese que

$$\left[\begin{array}{c|c} c & B^t \\ \hline B & A \end{array}\right]\left[\begin{array}{c|c} 1 & B^t \\ \hline O & -cI \end{array}\right] = \left[\begin{array}{c|c} c & O \\ \hline B & BB^t - cA \end{array}\right] = M'$$

Como el rango de esta última matriz M' es. obviamente, una unidad mayor que el rango de M_c y como, por otra parte, el rango de M' es igual al rango de M, pues se obtiene de multiplicar a M por una matriz regular, es:

$$\operatorname{rang} M_c = (\operatorname{rang} M') - 1 = (\operatorname{rang} M) - 1$$

EJERCICIO

Dados un punto $P(a, a, a)$ y la cuádrica

$$5x^2 - 2z^2 + 4xz - 3yz + 7y + 3 = 0$$

Hallar: 1) si P es de la cuádrica, el plano tangente en P; 2) si P no es de la cuádrica, el cono circunscrito desde P.

RESOLUCION

1. Según se acaba de obtener (véase [329], [2]), el plano tangente en P es:

$$[1 \;\; a \;\; a \;\; a]\begin{bmatrix} 3 & 0 & 7/2 & 0 \\ 0 & 5 & 0 & 2 \\ 7/2 & 0 & 0 & -3/2 \\ 0 & 2 & -3/2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ x \\ y \\ z \end{bmatrix} = O$$

o sea,

$$(7a)x + \frac{7-3a}{2}y - \frac{3a}{2}z + \frac{-12+7a}{2} = 0$$

2. El cono circunscrito es, según [329], [1]:

$$[7ax + \tfrac{1}{2}(7-3a)y - \tfrac{1}{2}3az + \tfrac{1}{2}(-12+7a)]^2 =$$
$$= (4a^2 + 7a + 3)(5x^2 - 2z^2 + 4xz - 3yz + 7y + 3)$$

PUNTOS ELIPTICOS, PARABOLICOS E HIPERBOLICOS

[330]

En el espacio afín tridimensional y respecto de una referencia cartesiana, se considera la cuádrica cuya matriz es M (simétrica 4×4). Sea P un punto cualquiera de la cuádrica, en el que se considera el plano tangente a ella (que la corta según dos rectas que pasan por P). Se dice que el punto P es *elíptico*, *parabólico* o *hiperbólico* de la cuádrica según que el plano tangente en P la corte según dos rectas imaginarias (esto es, un punto real), dos rectas confundidas (una recta doble) o dos rectas distintas (y reales), respectivamente.

Se verifica que todos los puntos de una misma cuádrica son de la misma naturaleza, esto es, todos elípticos, todos parabólicos o todos hiperbólicos. Más exactamente (*):

M = matriz de la cuádrica	Naturaleza de sus puntos
rang $M = 4$ y sig $M = (2, 2)$	todos hiperbólicos
rang $M = 4$ y sig $M \neq (2, 2)$ (**)	todos elípticos
rang $M = 3$ (cono o cilindro)	todos parabólicos (***)

(*) Sólo se consideran cuádricas ordinarias y conos o cilindros; los pares de planos carecen aquí de interés.

(**) Es decir, sig M valiendo $\{4, 0\}$ ó $\{3, 1\}$. El caso $\{4, 0\}$ es obvio, pues corresponde a cuádrica imaginaria.

(***) Excepto el vértice, en el que no hay un plano tangente.

DEMOSTRACION

Sabemos (ver [323]) que la signatura de $\pm M$ y, por tanto, el rango de M son invariantes frente al cambio de referencia; por ello estamos autorizados a elegir la referencia que más nos convenga. Tomemos unos nuevos ejes coordenados (coordenadas absolutas x_1, x_2, x_3) de manera que P sea el nuevo origen y el plano tangente a la cuádrica en P sea el $x_3 = 0$; sea

$$\sum_{i,j=1}^{3} a_{ij}x_i x_j + 2\sum_{h=1}^{3} b_h x_h + c = 0 \qquad [1]$$

la ecuación de la cuádrica en los nuevos ejes, a cuya matriz llamaremos M'. Como $P(0, 0, 0)$ es de la cuádrica, ha de ser $c = 0$. El plano tangente en el punto $P(0, 0, 0)$ es (según [329], [2]):

$$[1\ 0\ 0\ 0]M'\begin{bmatrix}1\\x_1\\x_2\\x_3\end{bmatrix} = O \qquad \text{o} \qquad b_1x_1 + b_2x_2 + b_3x_3 = 0$$

como este plano es el $x_3 = 0$, resulta que ha de ser $b_1 = b_2 = 0$ y $b_3 \neq 0$. Si dividimos la ecuación [1] por $b_3 \neq 0$, concluimos que, en la nueva referencia, la cuádrica tiene por matriz a la

$$M' = \begin{bmatrix}0 & 0 & 0 & 1\\0 & a'_{11} & a'_{12} & a'_{13}\\0 & a'_{21} & a'_{22} & a'_{23}\\1 & a'_{31} & a'_{32} & a'_{33}\end{bmatrix}$$

(para unos ciertos a'_{ij}). La intersección de la cuádrica con su plano tangente en $P(0, 0, 0)$ es, pues, la cónica que tiene por ecuaciones a las:

$$x_3 = 0 \quad \text{y} \quad a'_{11}x_1^2 + 2a'_{12}x_1x_2 + a'_{22}x_2^2 = 0$$

que, como cónica del plano $x_3 = 0$ (con las únicas coordenadas x_1 y x_2), admite por matriz a la

$$N' = \begin{bmatrix}0 & 0 & 0\\0 & a'_{11} & a'_{12}\\0 & a'_{21} & a'_{22}\end{bmatrix}$$

(Obsérvese que, realmente, se trata de un par de rectas). Cambiemos ahora los ejes Px_1 y Px_2 por otros Px'_1 y Px'_2, de su plano, en los que la matriz de la cónica sea diagonal, esto es, hagamos $x_1 = \alpha x'_1 + \beta x'_2$ y $x_2 = \gamma x'_1 + \delta x'_2$, donde (α, γ) y (β, δ) son columnas de una cualquiera de las matrices regulares que diagonalizan por congruencia a la matriz simétrica de los términos cuadráticos de N', o sea, tal que

$$\begin{bmatrix}\alpha & \gamma\\\beta & \delta\end{bmatrix}\begin{bmatrix}a'_{11} & a'_{12}\\a'_{21} & a'_{22}\end{bmatrix}\begin{bmatrix}\alpha & \beta\\\gamma & \delta\end{bmatrix} = \begin{bmatrix}a''_{11} & 0\\0 & a''_{22}\end{bmatrix}$$

(para unos ciertos a''_{11} y a''_{22}). En los nuevos ejes, las matrices de la cuádrica y del par de rectas en que ella corta al plano $x_3 = 0$ son, pues, del tipo:

$$M'' = \begin{bmatrix}0 & 0 & 0 & 1\\0 & a''_{11} & 0 & a'_{13}\\0 & 0 & a''_{22} & a'_{23}\\1 & a'_{31} & a'_{32} & a'_{33}\end{bmatrix} \quad \text{y} \quad N'' = \begin{bmatrix}0 & 0 & 0\\0 & a''_{11} & 0\\0 & 0 & a''_{22}\end{bmatrix}$$

Si se realizan operaciones elementales en las filas y en las columnas de M'', se obtiene fácilmente la siguiente matriz M''' que, por ser congruente con M'', tiene el mismo rango y la misma signatura que ella:

$$M''' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & a''_{11} & 0 & 0 \\ 0 & 0 & a''_{22} & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

A la vista de la forma que tienen las matrices M''' y N'', se deduce obviamente que

$$\text{sig}\, N'' = (p, q) \quad \Leftrightarrow \quad \text{sig}\, M''' = (p+1, q+1)$$

Como N'' es matriz del par de rectas que la cuádrica tiene en el plano $x_3 = 0$, de lo hasta aquí dicho se deduce que:

1. P es elíptico sii las rectas son imaginarias sii $\text{sig}\, N'' = \{2, 0\}$ (*) sii $\text{sig}\, M''' = \{3, 1\}$ (*) sii $\text{sig}\, M = \{3, 1\}$ (*) (el caso $\text{sig}\, M = \{4, 0\}$ no se presenta pues, para él, la cuádrica es imaginaria y, para empezar, no tendría ningún punto P en el que hallar el plano tangente).
2. P es hiperbólico sii las rectas son reales y distintas sii $\text{sig}\, N'' = (1, 1)$ sii $\text{sig}\, M''' = (2, 2)$ sii $\text{sig}\, M = (2, 2)$.
3. P es parabólico sii las rectas coinciden sii $\text{rang}\, N'' < 2$ sii $\text{rang}\, M''' < 4$ sii $\text{rang}\, M < 4$.

EJERCICIO

Determinar la naturaleza de los puntos de la siguiente cuádrica

$$3x^2 + 2y^2 + 9z^2 - 2xz + 8yz + 4x + 1 = 0$$

RESOLUCION

Diagonalizando por congruencia la matriz M de esta cuádrica, se llega fácilmente a la siguiente matriz D:

$$M = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 0 & -1 \\ 0 & 0 & 2 & 4 \\ 0 & -1 & 4 & 5 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

Como la signatura de D es (2, 2), también lo es la de M, luego sus puntos son, todos, hiperbólicos.

(*) Al poner que una signatura es $\{p, q\}$, se quiere señalar que puede valer (p, q) o (q, p).

CUADRICAS REGLADAS

[331]

Las cuádricas ordinarias[1] se dividen en dos clases (véase [320]): cuádricas con todos sus puntos elípticos, las cuales no contienen a ninguna recta, y cuádricas con todos sus puntos hiperbólicos, que se llaman *regladas* pues incluyen rectas y se caracterizan por la condición sig $M = (2, 2)$, donde M es matriz de la cuádrica.

En cualquier cuádrica reglada hay dos familias uniparamétricas de rectas, incluidas en la cuádrica, que se llaman *generatrices rectilíneas* de la cuádrica. Por cada punto de la cuádrica pasan dos generatrices rectilíneas, una de cada familia.

(1) Para las cuádricas degeneradas esta cuestión es obvia.

DEMOSTRACION

Antes de hablar de las cuádricas regladas, hablemos de las que no lo son. Si una cuádrica ordinaria es de las que tienen todos sus puntos elípticos, entonces no incluye a ninguna recta. En efecto: si incluyese a una recta, r, cortando a la cuádrica por un plano π que contenga a r, nos encontraríamos con que, como r forma parte de dicha intersección, ésta sería un par de rectas, r y s; por tanto, π sería el plano tangente en el punto P en el que se cortan r y s y P sería, entonces, un punto hiperbólico de la cuádrica, ya que el plano tangente a la cuádrica en P sería π.

Supongamos ya que la cuádrica es de las de puntos hiperbólicos y consideremos un punto P cualquiera de ella; el plano tangente en P a la cuádrica corta a ésta en dos rectas reales y distintas. Tomemos una referencia cartesiana $(P; \bar{e}_1, \bar{e}_2, \bar{e}_3)$ que tenga a dichas rectas como «eje de la x_1» y como «eje de la x_2». La ecuación de la cuádrica ha de verificarse, entonces, para todos los puntos de coordenadas $(x_1, 0, 0)$ y $(0, x_2, 0)$, por lo que no tendrá término independiente, ni términos en x_1, en x_2, en x_1^2 y en x_2^2, es decir, ha de ser del tipo

$$ax_3^2 + bx_1x_2 + cx_1x_3 + dx_2x_3 + ex_3 = 0 \qquad [1]$$

y, como el determinante de su matriz vale b^2e^2 y es no nulo, será $b \neq 0$ y $e \neq 0$; se puede pues suponer que es $b = 1$ (en caso contrario, divídase la ecuación por $b \neq 0$). Cambiemos la base $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$ de la referencia por la $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ siguiente:

$$\bar{u}_1 = \bar{e}_1, \quad \bar{u}_2 = \bar{e}_2 \quad \text{y} \quad \bar{u}_3 = -d\bar{e}_1 - c\bar{e}_2 + \bar{e}_3$$

con lo que las nuevas coordenadas x, y, z y las antiguas x_1, x_2, x_3 están relacionadas por:

$$x_1 = x - dz, \quad x_2 = y - cz, \quad x_3 = z \qquad [2]$$

Llevando estas expresiones [2] a la ecuación [1] de la cuádrica (no se olvide que $b = 1$), se obtiene la ecuación de ésta en los nuevos ejes; que es, pues, del tipo:

$$a'z^2 + xy + ez = 0 \quad (\text{donde } e \neq 0) \qquad [3]$$

Veamos ahora que por cada punto de cualquiera de los ejes Px o Py (que están incluidos en la cuádrica) pasa una generatriz rectilínea de la cuádrica, además del propio eje. Consideremos el eje Px (para el Py se procede de igual modo): tomemos un punto $(\alpha, 0, 0)$, cualquiera del eje Px, y busquemos una recta que pase por él, es decir, del tipo

$$x = \alpha + hz, \quad y = kz \qquad [4]$$

que forme parte de la cuádrica, esto es, tal que al llevar las expresiones [4] a la ecuación [3] nos quede una relación (en z) que se verifique para todo z; haciendo esto, se obtiene que ha de ser

$$(a' + hk)z^2 + (\alpha k + e)z = 0 \quad \text{para todo } z \in \mathbb{R}$$

luego ha de ser $k = -e/\alpha$ y $h = (a'\alpha)/e$, lo que nos da la generatriz:

$$x = \alpha + \frac{a'\alpha}{e} z, \quad y = -\frac{e}{\alpha} z \quad (\text{para cada } \alpha \neq 0) \qquad [5]$$

Análogamente se comprueba que, por cada punto $(0, \beta, 0)$, del eje Py, pasa la siguiente generatriz rectilínea de la cuádrica:

$$x = -\frac{e}{\beta} z, \quad y = \beta + \frac{a'\beta}{e} z \quad (\text{para cada } \beta \neq 0) \qquad [6]$$

Por cada punto de la cuádrica pasan dos generatrices rectilíneas, una de cada una de las familias [5] y [6]. Comprobémoslo para una de ellas, la [5], por ejemplo: dado un punto (x_0, y_0, z_0) de la cuádrica, que verificará a la ecuación [3], para que por él pase una generatriz de la familia [5], debe haber un $\alpha \neq 0$ tal que

$$\begin{cases} x_0 = \alpha + \dfrac{a'\alpha}{e} z_0 \\ y_0 = -\dfrac{e}{\alpha} z_0 \end{cases} \quad \text{o sea,} \quad \begin{cases} \alpha = \dfrac{ex_0}{e + a'z_0} \\ \alpha = \dfrac{-ez_0}{y_0} \end{cases}$$

lo que requiere que las dos expresiones de α sean iguales y esto es cierto, pues su igualación equivale, como se comprueba trivialmente, a:

$$a'z_0^2 + x_0y_0 + ez_0 = 0$$

que se verifica, pues es la expresión de que el punto (x_0, y_0, z_0) satisface a la ecuación [3] de la cuádrica.

EJERCICIO

Hallar las generatrices rectilíneas de la cuádrica reglada

$$xy - xz - z = 0$$

RESOLUCION

Buscamos las generatrices de entre las rectas de ecuación genérica:

$$x = \alpha + hz \quad , \quad y = \beta + kz \qquad [1]$$

Para que estas rectas estén incluidas en la cuádrica, al llevar los x e y que dan las expresiones [1] a la ecuación de la cuádrica, la ecuación de segundo grado (en z) que se obtiene ha de verificarse para todo z; esta ecuación es:

$$[h(k-1)]z^2 - [h\beta + k\alpha - \alpha - 1]z + [\alpha\beta] = 0$$

Para que esta ecuación se cumpla para todo $z \in \mathbb{R}$, han de ser nulos sus tres coeficientes, esto es, ha de ser:

$$h(k-1) = 0, \quad h\beta + k\alpha - \alpha - 1 = 0, \quad \alpha\beta = 0$$

Este sistema tiene dos soluciones (una para $\alpha = 0$ y la otra para $\beta = 0$):

$$\begin{cases} 1.^{a} \text{ solución: } \alpha = 0, \quad h = 1/\beta, \quad k = 1 \\ 2.^{a} \text{ solución: } \beta = 0, \quad k = (1+\alpha)/\alpha, \quad h = 0 \end{cases}$$

Por tanto, las dos familias de generatrices rectilíneas de la cuádrica son:

$$\begin{cases} x = \dfrac{z}{\beta} \\ y = \beta + z \end{cases} \quad (\beta \text{ parámetro}) \quad \text{y} \quad \begin{cases} x = \alpha \\ z = \dfrac{1+\alpha}{\alpha}\, z \end{cases} \quad (\alpha \text{ parámetro})$$

14.5. ANALISIS AFIN DE LAS CUADRICAS

Lo hasta aquí dicho acerca de las cuádricas en lo que va de capítulo han sido, casi exclusivamente, definiciones y propiedades de carácter proyectivo, esto es, aquellos que se conservan frente a los cambios de coordenadas homogéneas. A partir de ahora, entra otro elemento en juego; el plano del infinito como plano especial del espacio, plano distinto de los demás. Las concomitancias entre las cuádricas y los puntos impropios, o del infinito, nos va a ocupar en lo que sigue.

PUNTOS DEL INFINITO. CLASIFICACION AFIN

[332]

En el espacio tridimensional afín, se considera la cuádrica que, respecto de una cierta referencia cartesiana (coordenadas homogéneas X_0, X_1, X_2, X_3), admite por ecuación homogénea a:

$$\sum_{i,j=1}^{3} a_{ij}X_iX_j + 2\sum_{k=1}^{3} b_kX_0X_k + cX_0^2 = 0$$

Llamemos A a la matriz de términos cuadráticos; esto es:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Se llaman puntos del infinito de la cuádrica a los puntos del infinito que satisfacen a su ecuación homogénea; el lugar geométrico de estos puntos se llama *cónica del infinito* de la cuádrica y tiene las siguientes ecuaciones:

$$\sum_{i,j=1}^{3} a_{ij}X_iX_j = 0, \quad X_0 = 0 \qquad [1]$$

Atendiendo a su cónica del infinito, las cuádricas se clasifican como sigue:

La cuádrica se dice:	si su cónica del infinito es:	
de tipo ELIPTICO	Imaginaria (sin puntos)	$\Leftrightarrow \operatorname{sig} A = \{3, 0\}$ (**)
de tipo PARABOLICO	Degenerada (dos rectas)	$\Leftrightarrow \operatorname{rang} A < 3$
de tipo HIPERBOLICO	Ordinaria y real (*)	$\Leftrightarrow \operatorname{sig} A = \{2, 1\}$ (**)

Se dice que la cuádrica es, respectivamente, exterior, tangente o secante con el plano del infinito.

(*) Ordinaria y real significa no degenerada (no es un par de rectas) y con algún punto real.

(**) Sig $A = \{p, q\}$ significa que la signatura de A es (p, q) o (q, p).

DEMOSTRACION

Los puntos del infinito (véase [243] y [245]) son los puntos del espacio ampliado que tienen $X_0 = 0$. Al añadir esta condición a la ecuación de la cuádrica se obtienen las ecuaciones [1].

La cónica del infinito será imaginaria (carecerá de puntos reales) si, llamando X a la columna de coordenadas X_1, X_2, X_3, la relación $X^tAX = O$ no se satisface para ningún $X \neq O$, es decir, si A es definida, lo que acontece si su

signatura es (3, 0) o (0, 3). La cónica del infinito será degenerada si la expresión $\Sigma a_{ij}X_iX_j$ (para $i, j = 1, 2, 3$) se puede expresar como producto de dos expresiones lineales, lo que sucede si A es singular, es decir, si rang $A < 3$. Finalmente, la cónica del infinito es ordinaria y real si no es ni degenerada ni imaginaria, lo que significa que rang $A = 3$ y sig $A \neq \{3, 0\}$, que equivale a sig $A = \{2, 1\}$.

EJERCICIO

Analizar, en función de los valores del parámetro $\alpha \in \mathbb{R}$, si la siguiente cuádrica es de tipo elíptico, parabólico o hiperbólico

$$x^2 + 5y^2 + \alpha z^2 + 4xy - 2xz + 2yz + 9x - 3y + 7z - 2 = 0$$

RESOLUCION

Realizando operaciones elementales, en filas y columnas de la matriz A, de los términos cuadráticos de la cuádrica, se obtienen las siguientes matrices, congruentes con A y, por ello, con su misma signatura:

$$A\begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 1 \\ -1 & 1 & \alpha \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & \alpha - 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha - 10 \end{bmatrix} = D$$

Por tanto, en función de α, se tiene:

$\alpha > 10 \Rightarrow \text{sig}\, A = \text{sig}\, D = (3, 0) \Rightarrow$ cuádrica de tipo elíptico

$\alpha = 10 \Rightarrow \text{rang}\, A = \text{rang}\, D = 2 < 3 \Rightarrow$ cuádrica de tipo parabólico

$\alpha < 10 \Rightarrow \text{sig}\, A = \text{sig}\, D = (2, 1) \Rightarrow$ cuádrica de tipo hiperbólico

EJERCICIO

Se considera una cuádrica que, en cierta referencia cartesiana (coordenadas x, y, z), admite ecuación del tipo:

$$2\alpha yz + 2\beta zx + 2\gamma xy + 2ax + 2by + 2cz + h = 0,$$

donde α, β y γ no son los tres nulos.

Analícese si dicha cuádrica es de tipo elíptico, parabólico o hiperbólico.

RESOLUCION

La matriz de los términos cuadráticos de la cuádrica es:

$$A = \begin{bmatrix} 0 & \gamma & \beta \\ \gamma & 0 & \alpha \\ \beta & \alpha & 0 \end{bmatrix}$$

Como $\det A = 2\alpha\beta\gamma \neq 0$, la cuádrica no puede ser de tipo parabólico. Obsérvese, por otra parte, que los puntos del infinito de los ejes de coordenadas verifican a la ecuación de la cuádrica, luego ésta, por tener puntos en el infinito, no puede ser de tipo elíptico. Por no poder ser ni de tipo parabólico ni de tipo elíptico, es forzosamente de tipo hiperbólico.

ELIPSOIDES, PARABOLOIDES E HIPERBOLOIDES

[333]

Las cuádricas no degeneradas, atendiendo a su cónica del infinito, se llaman:

- ***Elipsoides*** **las de tipo elíptico (sin puntos en el infinito) y que además son ordinarias.**
- ***Paraboloides*** **las de tipo parabólico (tangente al plano del infinito) y que además son ordinarias.**
- ***Hiperboloides*** **las de tipo hiperbólico (secantes con el plano del infinito) y que además son ordinarias.**

La cuádrica que, respecto de una referencia cartesiana del espacio, tiene a M por matriz (simétrica; 4×4) y a A como submatriz de los términos cuadráticos, es elipsoide, es paraboloide o es hiperboloide si $\det M \neq 0$ y, respectivamente, es $\operatorname{sig} A = \{3, 0\}$, es $\operatorname{rang} A > 3$ o es $\operatorname{sig} A = \{2, 1\}$.

COMPROBACION

Los resultados que se acaban de expresar son consecuencias obvias de [327], I (la cuádrica es ordinaria si $\det M \neq 0$) y de [332] (clasificación afín de las cuádricas).

EJERCICIO

Hallar los valores de α y β para los que la siguiente cuádrica es una paraboloide:

$$x^2 - z^2 + 4xy + 2\alpha yz + 2x - \beta y + 1 = 0$$

RESOLUCION

Los determinantes de las matrices M, de la cuádrica, y A, de los términos cuadráticos valen:

$$\det A = 4 - \alpha^2 \quad \text{y} \quad \det M = 4 - \alpha^2 + \left(\frac{\beta}{2} - 2\right)^2$$

La cuádrica será de tipo parabólico si es $\det A = 0$, o sea, para $\alpha = \pm 2$. Para que sea paraboloide, ha de ser, además, $\det M \neq 0$, o sea $\beta \neq 4$.

EJERCICIO

Considérese un cilindro C (cuádrica con un solo punto singular, que está en el infinito); si M y A son la matriz de esta cuádrica y su submatriz de los términos cuadráticos, en una determinada referencia cartesiana, se sabe, pues, que $\operatorname{rang} M = 3$ y $\operatorname{rang} A < 3$. Córtese el cilindro por un plano π cualquiera que no sea paralelo a las generatrices del cilindro y considérese la cónica intersección $C \cap \pi$. Compruébese que:

$$
\begin{aligned}
&C \cap \pi \text{ elipse imaginaria} &&\Leftrightarrow \quad \operatorname{sig} M = \{3, 0\} \text{ y } \operatorname{sig} A = \{2, 0\} \\
&C \cap \pi \text{ elipse real} &&\Leftrightarrow \quad \operatorname{sig} M = \{2, 1\} \text{ y } \operatorname{sig} A = \{2, 0\} \\
&C \cap \pi \text{ hipérbola} &&\Leftrightarrow \quad \operatorname{sig} M = \{2, 1\} \text{ y } \operatorname{sig} A = \{1, 1\} \\
&C \cap \pi \text{ parábola} &&\Leftrightarrow \quad \operatorname{sig} M = \{2, 1\} \text{ y } \operatorname{sig} A = \{1, 0\}
\end{aligned}
$$

(decir que una signatura es $\{p, q\}$ significa que es (p, q) o (q, p)).

RESOLUCION

Recurramos a una nueva referencia cartesiana, $Ox_1x_2x_3$, en la que el plano Ox_1x_2 sea π y el eje Ox_3 tenga la dirección de las generatrices del cilindro, con lo que el punto singular del cilindro tiene por coordenadas homogéneas a las $(0, 0, 0, 1)$ y las matrices M' y A' (del cilindro y de sus términos cuadráticos) en la nueva referencia son, por tanto, del tipo:

$$
M' = \left[\begin{array}{cccc} c & b_1 & b_2 & 0 \\ b_1 & a_{11} & a_{12} & 0 \\ b_2 & a_{21} & a_{22} & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right] = \begin{bmatrix} N & O \\ O & 0 \end{bmatrix} \qquad A' = \left[\begin{array}{ccc} a_{11} & a_{12} & 0 \\ a_{12} & a_{22} & 0 \\ \hline 0 & 0 & 0 \end{array}\right] = \begin{bmatrix} B & O \\ O & 0 \end{bmatrix}
$$

Nótese que M' y N tienen, obviamente, la misma signatura y que lo mismo acontece con A' y B. Por otra parte, según ya sabemos (véase [323]), M' y A' tienen las mismas signaturas que M y A o que $-M$ y $-A$. Por tanto, se concluye que

$$
[\operatorname{sig} M = \{p, q\}] \quad \Leftrightarrow \quad [\operatorname{sig} N = \{p, q\}]
$$

y

$$
[\operatorname{sig} A = \{p, q\}] \quad \Leftrightarrow \quad [\operatorname{sig} B = \{p, q\}]
$$

Al expresar la cónica $C \cap \pi$ en los ejes Ox_1x_2, de su plano, resulta evidente que

su matriz es N y la submatriz de los términos cuadráticos es B. Teniendo ahora en cuenta lo dicho en la observación que sigue a [312], se concluye que:

$C \cap \pi$ es una:	sig N y sig B	sig M y sig A
Elipse imaginaria	$\Leftrightarrow$ {3, 0} y {2, 0}	$\Leftrightarrow$ {3, 0} y {2, 0}
Elipse real	$\Leftrightarrow$ {2, 1} y {2, 0}	$\Leftrightarrow$ {2, 1} y {2, 0}
Hipérbola	$\Leftrightarrow$ {2, 1} y {1, 1}	$\Leftrightarrow$ {2, 1} y {1, 1}
Parábola	$\Leftrightarrow$ {2, 1} y {1, 0}	$\Leftrightarrow$ {2, 1} y {1, 0}

CLASIFICACION GENERAL DE LAS CUADRICAS

[334]

En el siguiente cuadro se resume la clasificación de las cuádricas [1]; en él, M es la matriz de la cuádrica y A es su submatriz de los términos cuadráticos (en una referencia cartesiana cualquiera):

rang M	rang A	sig M [3]	sig A [3]	Tipo de cuádrica
rang $M = 4$ (cuádricas ordinarias)	rang $A = 3$ (con centro) [2]	{4, 0}	(3, 0) (*)	ELIPSOIDE IMAGINARIO
		{3, 1}	(3, 0]	ELIPSOIDE REAL
		{3, 1}	(2, 1]	HIPERBOLOIDE NO REGLADO
		{2, 2}	(2, 1] (*)	HIPERBOLOIDE REGLADO
	rang $A = 2$ (sin centro) [2]	{3, 1}	(2, 0] (*)	PARABOLOIDE NO REGLADO
		{2, 2}	(1, 1] (*)	PARABOLOIDE REGLADO
rang $M = 3$ (con un punto singular)	rang $A = 3$ (conos)	{3, 0}	(3, 0] (*)	CONO IMAGINARIO (con vértice real)
		{2, 1}	(2, 1] (*)	CONO REAL
	(cilindros) rang $A = 2$	{3, 0}	(2, 0] (*)	CILINDRO IMAGINARIO
		{2, 1}	(2, 0]	CILINDRO ELIPTICO (real)
		{2, 1}	(1, 1]	CILINDRO HIPERBOLICO
	rang $A = 1$	{2, 1}	(1, 0]	CILINDRO PARABOLICO

(1) Se excluyen los pares de planos: tienen escaso interés y complican en exceso la clasificación.
(2) La definición de centro se dará más adelante (véase [335]).
(3) Decir que una signatura es $\{p, q\}$ significa que vale (p, q) o (q, p).
(*) Esta exigencia es innecesaria (superabundante), por ser consecuencia de las anteriores.

COMPROBACION

Lo que se acaba de recoger en el cuadro precedente es consecuencia de lo dicho hasta ahora: 1) Al hablar de cuádricas ordinarias y degeneradas y su clasificación en función de sus puntos singulares (véase [327]). 2) Al hablar de cuádricas reales e imaginarias (véase [328]. 3) Al hablar de cuádricas regladas y no regladas (véase [331]). 4) Al hablar de los puntos del infinito de las cuádricas (véase [332]). 5) Al hablar de los elipsoides, los hiperboloides y los paraboloides (véase [333]). 6) Al hablar de los cilindros en el último de los ejercicios que hemos realizado. Sin embargo, se hacen necesarias algunos comentarios y aclaraciones.

Se llaman elipsoides imaginarios a las cuádricas imaginarias (ordinarias), ya que son de tipo elíptico (sin puntos en el infinito) y ordinarias. Salvo los elipsoides imaginarios, que tienen sig $M = \{4, 0\}$, las demás cuádricas ordinarias tienen la sig M valiendo $\{2, 2\}$ ó $\{3, 1\}$ según que, respectivamente sean regladas o no regladas. Para comprobar que las condiciones que se han marcado con un asterisco (*), en el cuadro, son consecuencia de la condición que figura en el correspondiente recuadro de su izquierda, se puede, en los casos en que ello no sea evidente, tomar una referencia cartesiana adecuada al caso. Para los cilindros, basta con recordar lo dicho en el último ejercicio, el de la página 659, y tener en cuenta que el cilindro se llama imaginario si carece de puntos reales propios (aunque tiene un punto real en el infinito) y se dice que es elíptico, hiperbólico o parabólico según que, al cortarle por un plano que no sea paralelo a sus generatrices, se obtiene una elipse, una hipérbola o una parábola, respectivamente.

EJERCICIO

Clasificar la siguiente cuádrica, en función de los valores del parámetro $p \in \mathbb{R}$:

$$2x^2 - y^2 + pz^2 - 2xz + 2yz - 4y - 2 = 0$$

RESOLUCION

Realizando operaciones elementales, en las filas y en las columnas, se obtienen fácilmente matrices diagonales congruentes con las matrices M y A, de la cuádrica y de sus términos cuadráticos; las diagonales de estas matrices resultan ser:

Diagonal correspondiente a A: $(2, -1, \alpha + 1/2)$

Diagonal correspondiente a M: $(-2, 2, 1, 4\alpha - 6)$

Por tanto, de acuerdo con el cuadro de [334], es:

	sig M	sig A	Tipo de cuádrica
$\alpha < -1/2$	$\{2, 2\}$	$\{2, 1\}$	Hiperboloide reglado
$\alpha = -1/2$	$\{2, 2\}$	$\{1, 1\}$	Paraboloide reglado
$-1/2 < \alpha < 3/2$	$\{2, 2\}$	$\{2, 1\}$	Hiperboloide reglado
$\alpha = 3/2$	$\{2, 1\}$	$\{2, 1\}$	Cono real
$\alpha > 3/2$	$\{3, 1\}$	$\{2, 1\}$	Hiperboloide no reglado

CUADRICAS CON CENTRO

[335]

En el espacio afín tridimensional y respecto de una cierta referencia cartesiana (coordenadas x_1, x_2, x_3), se considera la cuádrica que admite por ecuación a:

$$\sum_{i,j=1}^{3} a_{ij}x_ix_j + 2\sum_{h=1}^{3} b_hx_h + c = 0;$$

cuya matriz es

$$M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \quad \text{donde} \quad \begin{cases} A = [a_{ij}] \text{ es la matriz de los términos cuadráticos} \\ B = [b_h] \text{ (matriz columna } 3\times 1) \end{cases}$$

Se dice que un punto P_C es *centro* de la cuádrica si es centro de simetría de la misma. A este respecto, se verifica que:

I. Si un punto $P_C(x_1, x_2, x_3)$ es solución del sistema de ecuaciones:

$$\sum_{j=1}^{3} a_{ij}x_j = -b_i \quad (\text{para } i = 1, 2, 3) \quad \text{o} \quad AX = -B \qquad [1]$$

entonces P_C es centro de la cuádrica. En relación con el centro, de la discusión de [1] se obtiene:

1. Si $\delta_0 = \det A \neq 0$, entonces la cónica tiene un centro, que es:

$$P_C\left(\frac{\delta_1}{\delta_0}, \frac{\delta_2}{\delta_0}, \frac{\delta_3}{\delta_0}\right) \qquad [2]$$

donde, δ_0, δ_1, δ_2, δ_3 son los adjuntos de los elementos de la primera fila de M. Los elipsoides, hiperboloides y conos tienen centro (para estos últimos, el centro es su punto singular).

2. Los paraboloides (que tienen $\delta_0 = \det A = 0$) carecen de centro (*).

II. Si una cuádrica tiene centro, al que llamaremos $P_C(\alpha_1, \alpha_2, \alpha_3)$, y se traslada la referencia cartesiana para que tenga su origen en P_C (nuevas coordenadas $x_i' = x_i - \alpha_i$, para $i = 1, 2, 3$), entonces la ecuación de la cuádrica toma la siguiente forma (que carece de términos lineales):

$$\sum_{i,j=1}^{3} a_{ij}x_i'x_j' + c' = 0 \quad \text{donde} \quad c' = c + \sum_{h=1}^{3} b_h\alpha_h \qquad [3]$$

(*) Un cilindro (rang $M = 3$ y rang $A = 2$) tiene infinitos centros, que son soluciones de [1]. Este sistema tiene, ahora, infinitas soluciones pues rang A = rang $\overline{AB} < 3$.

DEMOSTRACION

La demostración de [313], con los retoques pertinentes, es también de aplicación a este caso: sin más que extender los sumatorios de 1 a 3, en lugar de hacerlo de 1 a 2, considerando tres coordenadas en lugar de dos, teniendo en cuenta que ahora M y A son de tamaños 4×4 y 3×3, en vez de serlo de 3×3 y 2×2, hablando de elipsoides, hiperboloides y paraboloides, en lugar de hacerlo de elipses, hipérbolas y parábolas, y con algún otro retoque similar, la demostración del [313] se convierte en la demostración de las propiedades incluidas en este apartado [335].

EJERCICIO

[336] En el espacio afín tridimensional y respecto de una cierta referencia cartesiana rectangular, se considera una cuádrica; supongamos que su ecuación es:

$$\sum_{i,j=1}^{3} a_{ij}x_i x_j + 2\sum_{h=1}^{3} b_h x_h + c = 0$$

cuya matriz es

$$M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix}$$

Si la cuádrica tiene centro P_C (se supone $\det A \neq 0$) y llamando X_C a la columna de coordenadas del centro, pruébese que:

1. $$M\begin{bmatrix} 1 \\ X_C \end{bmatrix} = \begin{bmatrix} \alpha \\ O \end{bmatrix} \quad \text{donde} \quad \alpha = \frac{\det M}{\det A} \qquad [1]$$

2. El plano polar del centro (véase [329], III) es el plano del infinito (aquí se supone que la cuádrica es ordinaria, $\det M \neq 0$).

RESOLUCION

1. Como $AX_C = -B$ (véase [335], [1]), multiplicando por bloques, se tiene:

$$M\begin{bmatrix} 1 \\ X_C \end{bmatrix} = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix}\begin{bmatrix} 1 \\ X_C \end{bmatrix} = \begin{bmatrix} c + B^t X_C \\ B + AX_C \end{bmatrix} = \begin{bmatrix} \alpha \\ O \end{bmatrix} \quad \text{con} \quad \alpha = c + B^t X_C$$

Ahora bien, como las coordenadas del centro son (según [335], I, 1) los números δ_i/δ_0 (para $i = 1, 2, 3$), donde los δ_0, δ_1, δ_2, δ_3 son los adjuntos de los elementos de la primera fila de M, resulta que

$$\alpha = c + \frac{1}{\delta_0}\sum_{h=1}^{3} b_h\delta_h = \frac{1}{\delta_0}(c\delta_0 + b_1\delta_1 + b_2\delta_2 + b_3\delta_3) = \frac{\det M}{\det A}$$

ya que el paréntesis anterior es el desarrollo del det M por los elementos de su primera fila y $\delta_0 = \det A$.

2. Llamando X a la columna de coordenadas homogéneas (X_0, X_1, X_2, X_3) de un punto del plano buscado, esto es, del plano polar del centro, dicho plano polar admitirá por ecuación (según [329], III) a $[1\ X_C^t]MX = O$; trasponiendo esta ecuación y recurriendo a la igualdad [1], se tiene que

$$O = X^t M \begin{bmatrix} 1 \\ X_C \end{bmatrix} = X^t \begin{bmatrix} \alpha \\ O \end{bmatrix} = [X_0 \alpha]$$

Como $\alpha \neq 0$ (pues det $M \neq 0$), la ecuación del plano polar del centro es la $X_0 = 0$, es decir, dicho plano es el del infinito, como había que comprobar.

EJERCICIO

Dada la cuádrica de ecuación (a y b parámetros).

$$x^2 + y^2 + az^2 + 4xy + 6yz - 2x + 8y - 2bz + 1 = 0$$

hallar su entro (o sus centros), discutiendo la existencia de los mismos en relación con el tipo de cuádrica.

RESOLUCION

La matriz M de la cuádrica y la submatriz A de los términos cuadráticos son:

$$M = \begin{bmatrix} 1 & -1 & 4 & -b \\ -1 & 1 & 2 & 0 \\ 4 & 2 & 1 & 3 \\ -b & 0 & 3 & a \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 0 & 3 & a \end{bmatrix}$$

Luego el sistema que proporciona los posibles centros (en función de a y b) es el [1], que se da a continuación; realizando, en él, operaciones elementales se obtiene el [2]:

$$\left.\begin{array}{l} -1 + x + 2y = 0 \\ 4 + 2x + y + 3z = 0 \\ -b + 3y + az = 0 \end{array}\right\} \ [1] \qquad \left.\begin{array}{l} x + 2y = 1 \\ -y + z = -2 \\ (a+3)z = b - 6 \end{array}\right\} \ [2]$$

La discusión de [2] nos conduce a:

- Si $a \neq -3$ la cuádrica tiene un (y sólo un) centro. Nótese que para $a \neq -3$ es det $A \neq 0$ y la cuádrica es de tipo elíptico o hiperbólico (no parabólico). En este caso, el centro es, pues, el punto de coordenadas

$$x = \frac{3 - 3a - 2b}{a+3} \quad , \quad y = \frac{2a + b}{a+3} \quad , \quad z = \frac{b - c}{a+3}$$

- Si $a=-3$ y $b\neq 6$, entonces la cuádrica carece de centro. Nótese que para $a=-3$ y $b\neq 6$ es $\det A=0$ y $\det M=3(b-6)^2\neq 0$, luego la cuádrica es un paraboloide.
- Si $a=-3$ y $b=6$, entonces la cuádrica tiene infinitos centros que son los puntos $(-3-2z,\ 2+z,\ z)$ para $z\in\mathbb{R}$. Nótese que ahora es $\operatorname{rang} M=3$ y $\operatorname{rang} A=2$, con lo que la cuádrica es un cilindro.

DIRECCION DE UN PARABOLOIDE

[337]

En el espacio afín tridimensional y respecto de una cierta referencia cartesiana, se considera el paraboloide cuya matriz es M; sea A su submatriz de términos cuadráticos. Llamemos δ_0, δ_1, δ_2, δ_3 a los adjuntos de los elementos de la primera fila de M (nótese que $\delta_0=\det A=0$, pues se habla de un paraboloide). Entonces se verifica que:

1. El punto, P_∞, de tangencia del paraboloide con el plano del infinito (*) tiene por coordenadas homogéneas a las $P_\infty(0,\ \delta_1,\ \delta_2,\ \delta_3)$; esto es, P_∞ es el punto del infinito de la dirección del vector $\bar{\delta}(\delta_1,\ \delta_2,\ \delta_3)$, a la que se llama *dirección del paraboloide*.
2. El vector $\bar{\delta}(\delta_1,\ \delta_2,\ \delta_3)$ es un vector propio de la matriz A, correspondiente a su autovalor $\lambda=0$.

NOTA: Todo lo anterior es igualmente válido para los cilindros. En este caso, la dirección de $\bar{\delta}$ es la de las generatrices del cilindro.

(*) P_∞ es el punto singular de la cónica del infinito de la cuádrica.

DEMOSTRACION

Antes de ninguna otra cosa, nótese que, como $\det A=0$, la matriz A admite el autovalor $\lambda=0$. Obsérvese también que, llamando

$$\Delta^t=[\delta_0\ \delta_1\ \delta_2\ \delta_3]=[0\ \delta_1\ \delta_2\ \delta_3]$$

es:

$$\Delta^t M=[\det M\ 0\ 0\ 0] \qquad [1]$$

ya que la suma de los productos de los elementos δ_0, δ_1, δ_2, δ_3 (los adjuntos de los elementos de la primera columna de M) por los respectivos elementos de la columna i-ésima de M vale: $\det M$ si $i=1$ y 0 si $i\neq 1$.

1. De [1] se deduce obviamente que $\Delta^t M\Delta=O$, lo que significa que el punto $P_\infty(0,\ \delta_1,\ \delta_2,\ \delta_3)$ pertenece a la cuádrica. Además, el plano tangente a la cuádrica en P_∞ es (véase [329], [2]) el $\Delta^t MX=O$ (donde X es la columna de las coordenadas homogéneas X_0, X_1, X_2, X_3 de un punto genérico del plano tangente); a la vista de [1], resulta que el plano tangente en P_∞ es el

$(\det M)X_0 = 0$, o sea el $X_0 = 0$, que es el del infinito, como había que comprobar.

2. Expresando la matriz M en términos de sus bloques usuales (en función de A, submatriz de los términos cuadráticos, de B, matriz de los términos lineales, y de c, término independiente), la igualdad [1] se puede poner en la forma:

$$\begin{bmatrix} \det M \\ 0 \\ 0 \\ 0 \end{bmatrix} = M\Delta = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \begin{bmatrix} 0 \\ \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}$$

De esta igualdad, multiplicando por bloques, se desprende que

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = A \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}$$

es decir, que el vector $\bar{\delta}(\delta_1, \delta_2, \delta_3)$ es vector propio de A correspondiente a su autovalor $\lambda = 0$.

EJERCICIO

[338] Dado un paraboloide, considérese unos ejes cartesianos, $Px_1x_2x_3$, de entre los muchos que cumplen las siguientes condiciones:

- El origen P es un punto (cualquiera) del paraboloide.
- El eje Px_3 es el que tiene la dirección del paraboloide.
- El plano Px_1x_2 es el plano tangente en P al paraboloide.

Pruébese que, en dicha referencia, la ecuación del paraboloide es del siguiente tipo (para ciertos $\alpha, \beta, \gamma \in \mathbb{R}$):

$$2x_3 = \alpha x_1^2 + 2\beta x_1 x_2 + \gamma x_2^2 \quad (\text{con } \alpha\gamma = \beta^2)$$

RESOLUCION

Denotemos a las matrices M y A (simétricas), de la cuádrica y de sus términos cuadráticos, del modo usual, esto es:

$$M = \begin{bmatrix} c & b_1 & b_2 & b_3 \\ b_1 & a_{11} & a_{12} & a_{13} \\ b_2 & a_{21} & a_{22} & a_{23} \\ b_3 & a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Como la dirección del paraboloide es la del vector $\bar{\delta}(0, 0, 1)$, éste es vector propio de A correspondiente al autovalor $\lambda = 0$; imponiendo que así sea, se obtiene que $a_{13} = a_{31} = a_{23} = a_{32} = a_{33} = 0$, con lo que la ecuación del paraboloide queda:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + 2b_3x_3 + c = 0 \qquad [1]$$

Como el paraboloide pasa por el origen, $P(0, 0, 0)$, ha de ser $c = 0$. Al cortar el paraboloide por el eje Px_1 (esto es, al hacer $x_2 = x_3 = 0$ en [1]) se ha de obtener la solución doble $x_1 = 0$, luego debe ser $2b_1 = 0$; procediendo de igual modo con el eje Px_2, se concluye que $2b_2 = 0$. Los únicos coeficientes no nulos que nos quedan son, pues, a_{11}, $2a_{12}$, a_{22} y $2b_3$. Como $\det M = -b_3^2(a_{11}a_{22} - a_{12}^2)$ y, por tratarse de un paraboloide, es $\det M \neq 0$, resulta que $b_3 \neq 0$. Dividiendo la ecuación [1] por b_3 y llamando

$$\alpha = a_{11}/b_3 \quad , \quad \beta = a_{12}/b_3 \quad , \quad \gamma = a_{22}/b_3$$

se obtiene la ecuación que se buscaba, en la que, como $\det M = \beta^2 - \alpha\gamma$ y es $\det M \neq 0$, se verifica que $\alpha\gamma \neq \beta^2$.

CONO ASINTOTICO DE UN HIPERBOLOIDE

[339]

En el espacio afín tridimensional y respecto de una referencia cartesiana (coordenadas x_1, x_2, x_3), se considera un hiperboloide; supongamos que su ecuación es:

$$\sum_{i,j=1}^{3} a_{ij}x_ix_j + 2b \sum_{k=1}^{3} b_kx_k + c = 0$$

cuya matriz es

$$M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} \quad \text{donde} \begin{cases} A = [a_{ij}] \text{ es la matriz de los términos cuadráticos} \\ B = [b_k] \text{ (matriz columna } 3 \times 1) \end{cases}$$

Se llama *cono asintótico* del hiperboloide al cono cincunscrito (véase [329]) al mismo desde su centro. La tangencia del cono asintótico con el hiperboloide se produce a lo largo de la cónica del infinito de éste; el plano polar del centro, respecto del hiperboloide, es el plano del infinito.

El cono asintótico tiene por ecuación a la

$$\sum_{i,j=1}^{3} a_{ij}x_ix_j + 2\sum_{k=1}^{3} b_kx_k = \alpha \quad \text{siendo} \quad \alpha = \frac{\det M}{\det A}$$

DEMOSTRACION

Según ya sabemos (véase el Ejercicio de [336], 2), el plano polar del centro es el plano del infinito y, por ello, la tangencia del cono asintótico (esto es, del

cono circunscrito desde el centro) se produce a lo largo de la cónica del infinito del hiperboloide (véase [329], III). La ecuación del cono asintótico es (según [329], I):

$$\left([1\ X_C^t]M\begin{bmatrix}1\\X\end{bmatrix}\right)^2=\left([1\ X_C^t]M\begin{bmatrix}1\\X_C\end{bmatrix}\right)\left([1\ X^t]M\begin{bmatrix}1\\X\end{bmatrix}\right) \qquad [1]$$

donde X_C es la columna de coordenadas absolutas del centro y X es la columna de coordenadas absolutas de un punto genérico del cono asintótico. Recuérdese que (según se vio en el ejercicio [336], 1) es:

$$M\begin{bmatrix}1\\X_C\end{bmatrix}=\begin{bmatrix}\alpha\\O\end{bmatrix} \quad \text{con} \quad \alpha=\frac{\det M}{\det A}$$

Llevando este resultado a la ecuación [1], se obtiene:

$$\left([\alpha\ O]\begin{bmatrix}1\\X\end{bmatrix}\right)^2=\left([\alpha\ O]\begin{bmatrix}1\\X_C\end{bmatrix}\right)\left([1\ X^t]M\begin{bmatrix}1\\X\end{bmatrix}\right)$$

Como el primer miembro vale α^2 y el primer paréntesis del segundo es α, resulta que la ecuación del cono asintótico es:

$$[1\ X^t]M\begin{bmatrix}1\\X\end{bmatrix}=\alpha \quad \text{o sea} \quad \sum_{i,j=1}^{3} a_{ij}x_ix_j+2\sum_{h=1}^{3} b_hx_h=\alpha$$

como había que comprobar.

EJERCICIO

Dada la cuádrica

$$x^2+3z^2+4xy+2z+2=0$$

compruébese que es un hiperboloide y hállense: 1) su cono asintótico; y 2) el cono que proyecta su cónica del infinito desde el origen.

RESOLUCION

Como $\det M=-20\neq 0$ y $\det A=-12<0$ (M y A son la matriz de la cuádrica y su submatriz de términos cuadráticos), la cuádrica es un hiperboloide (nótese que por ser $\det A<0$, es $\operatorname{sig} A=\{2, 1\}$).

1. Según lo dicho en [339], como ahora $\alpha=-20/12=-5/3$, el cono asintótico es:

$$x^2+3z^2+4xy+2z=-5/3$$

2. Como la cónica del infinito del hiperboloide es (en coordenadas homogéneas):

$$\begin{cases} x^2 + 3z^2 + 4xy = 0 \\ t = 0 \end{cases}$$

al proyectar esta cónica desde el origen se obtiene el cono:

$$x^2 + 3z^2 + 4xy = 0$$

14.6. ANALISIS METRICO DE LAS CUADRICAS

Vamos ahora a abordar aquellos aspectos de las cuádricas que dependen de las distancias y de los ángulos. Para que las expresiones que conducen a ellos no sean excesivamente complicadas, utilizaremos, de aquí en adelante, referencias cartesianas rectangulares, esto es, con la base de vectores ortonormal.

CAMBIO DE COORDENADAS RECTANGULARES

[340]

En el espacio afín euclídeo tridimensional, se considera una cuádrica cualquiera. Supongamos que, en una cierta referencia rectangular, M es la matriz de la cuádrica y que A es su submatriz de los términos cuadráticos. Supongamos también que, en otra referencia rectangular, las referidas matrices son M' y A'.

Si Q^* es la matriz del correspondiente cambio de coordenadas (véase [322]); esto es, si:

$$Q^* = \begin{bmatrix} 1 & O \\ C & Q \end{bmatrix} \text{ donde } \begin{cases} Q = \text{matriz ortogonal correspondiente al cambio de la base de la referencia rectangular} \\ C = \text{columna de coordenadas del nuevo origen} \end{cases}$$

$$\text{Entonces } \begin{cases} M' = Q^{*t} M Q^* & (M \text{ y } M' \text{ son matrices congruentes}) \\ A' = Q^t A Q & (A \text{ y } A' \text{ son ortogonalmente congruentes}) \end{cases}$$

COMPROBACION

Vale lo dicho en la comprobación de [316] (póngase ahora [322] donde allí se dice [303]).

EJERCICIO

Se considera la cuádrica que, en cierta referencia rectangular $(O;\ \bar{e}_1,\ \bar{e}_2,\ \bar{e}_3)$, tiene por ecuación a

$$3x^2 + 4xy + 3y^2 + z^2 + 2x - 2y + 1 = 0$$

Expresar dicha cuádrica en la referencia que tiene su origen en $P(-1,\ 1,\ 0)$ y que tiene por base a la que forman los vectores

$$\bar{u}_1 = \frac{\bar{e}_1}{2} - \frac{\bar{e}_2}{2} + \frac{\sqrt{2}}{2}\bar{e}_3 \quad , \quad \bar{u}_2 = \frac{-\bar{e}_1}{2} + \frac{\bar{e}_2}{2} + \frac{\sqrt{2}}{2}\bar{e}_3 \quad \text{y} \quad \bar{u}_3 = \frac{\sqrt{3}}{2}\bar{e}_1 + \frac{\sqrt{2}}{2}\bar{e}_2$$

A la vista del resultado que se obtenga, localizar ejes de simetría (ortogonal) de la cuádrica.

RESOLUCION

La base $(\bar{u}_1,\ \bar{u}_2,\ \bar{u}_3)$ es ortonormal, por lo que la nueva referencia también es rectangular. La matriz M de la cuádrica y la matriz Q^* del cambio de coordenadas son:

$$M = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 1 & 3 & 2 & 0 \\ -1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad Q^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1/2 & -1/2 & \sqrt{2}/2 \\ 1 & -1/2 & 1/2 & \sqrt{2}/2 \\ 0 & \sqrt{2}/2 & \sqrt{2}/2 & 0 \end{bmatrix}$$

Por tanto, la matriz M', en los nuevos ejes, es:

$$M' = Q^{*t} M Q^* = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

y la ecuación de la cuádrica en ellos es

$$x'^2 + y'^2 + 5z'^2 = 1$$

Esta cuádrica es un elipsoide cuyos ejes (de simetría ortogonal) son los Px', Py' y Pz', esto es las rectas (ortogonales dos a dos):

$$(x,\ y,\ z) = (-1,\ 1,\ 0) + \lambda(1,\ -1,\ \sqrt{2})$$
$$(x,\ y,\ z) = (-1,\ 1,\ 0) + \lambda(-1,\ 1,\ \sqrt{2})$$
$$(x,\ y,\ z) = (-1,\ 1,\ 0) + \lambda(1,\ 1,\ 0)$$

INVARIANTES METRICOS

[341]

En el espacio afín euclídeo tridimensional y respecto de una cierta referencia rectangular, se considera la cuádrica cuya matriz es M; sea $A = [a_{ij}]$ su submatriz de los términos cuadráticos y llamemos λ_1, λ_2 y λ_3 a los autovalores de A. Supongamos que, en una nueva referencia rectangular, los anteriores elementos son ahora M', $A' = [a'_{ij}]$, λ'_1, λ'_2 y λ'_3, respectivamente. Se verifica, entonces, que existe un cierto $\rho \in \mathbb{R}$ tal que:

1. $\det M' = \rho^4 (\det M)$.
2. $\det A' = \rho^3 (\det A)$ (*).
3. $a'_{11} + a'_{22} + a'_{33} = \rho(a_{11} + a_{22} + a_{33})$ (*).
4. $\lambda'_i = \rho\lambda_i$ (para $i = 1, 2, 3$).

Por ello, para señalar este resultado, se dice que $\det M$, $\det A$, $a_{11}+a_{22}+a_{33}$ (traza de A), λ_1, λ_2 y λ_3 son invariantes métricos de las cuádricas.

(*) Estas dos igualdades (2 y 3) son consecuencia de la última (4). Hay otro invariante (de menor interés) que junto con los 2 y 3 forma una terna equivalente a los tres invariantes de 4; este nuevo invariante es el: $\alpha'_{11} + \alpha'_{22} + \alpha'_{33} = \rho^2(\alpha_{11} + \alpha_{22} + \alpha_{33})$ (α_{ii} es el adjunto de a_{ii} en A).

DEMOSTRACION

Haciendo la oportuna adaptación, aquí es de aplicación lo dicho en la demostración de [317]. Donde allí se dice cónica, plano, dimensión $n = 2$ e $i = 1, 2$, debe aquí decirse cuádrica, espacio, dimensión $n = 3$ e $i = 1, 2, 3$.

Solamente hay que observar que ahora las relaciones 2 y 3 no equivalen, como en [317], a las relaciones 4, sino que son consecuencia de éstas. El motivo es que la ecuación característica tiene ahora 3 coeficientes, en lugar de 2; las relaciones 4 equivalen a la invarianza de dichos tres coeficientes y no sólo a la de dos de ellos. Aunque no tiene excesivo interés, no viene mal conocer que la invarianza de dicho tercer coeficiente conduce a la relación que figura en la nota (*) al pie del enunciado.

EJERCICIO

Se considera la cuádrica que, respecto de ciertos ejes rectangulares, admite por ecuación a

$$2x^2 + y^2 + z^2 + 4yz + 4x - 2y + 4z + 8 = 0$$

Esta cuádrica que, como se comprueba fácilmente, es un hiperboloide, tiene, por tanto, ecuación reducida del tipo

$$\alpha x'^2 + \beta y'^2 + \gamma z'^2 = 1$$

(x', y' y z' son las nuevas coordenadas rectangulares, pues las ecuaciones reducidas, según sabemos del capítulo anterior, lo son en ejes rectangulares). Hallar dicha ecuación reducida.

RESOLUCION

La cuádrica tiene por matriz, M, y por submatriz de los términos cuadráticos, A, a las siguientes:

$$M = \begin{bmatrix} c & B^t \\ B & A \end{bmatrix} = \left[\begin{array}{c|ccc} 8 & 2 & 2 & 0 \\ \hline 2 & 0 & -1 & 0 \\ -1 & 1 & 2 & -2 \\ -2 & 0 & 2 & 1 \end{array}\right]$$

Por tanto, los invariantes métricos de esta cuádrica ($\det M$ y los tres autovalores de A) son:

$$\det M = -30; \quad \lambda_1 = -1, \quad \lambda_2 = 2, \quad \lambda_3 = 3$$

Como sabemos que su ecuación reducida es del tipo:

$$\alpha x'^2 + \beta y'^2 + \gamma z'^2 = 1$$

resulta que, respecto de sus ejes, la nueva matriz M' de la cuádrica y la correspondiente submatriz de los términos cuadráticos, A', son

$$M' = \begin{bmatrix} c' & B'^t \\ B' & A' \end{bmatrix} = \begin{bmatrix} -1 & & & O \\ & \alpha & & \\ & & \beta & \\ O & & & \gamma \end{bmatrix}$$

y $\det M'$ y los autovalores de A' valen:

$$\det M' = -\alpha\beta\gamma \;\; ; \;\; \lambda_1' = \alpha \;\; , \;\; \lambda_2' = \beta \;\; , \;\; \lambda_3' = \gamma$$

Por tanto, según acabamos de comprobar (véase [341]), existe $\rho \in \mathbb{R}$ tal que

$$\left.\begin{array}{r} -\alpha\beta\gamma = -30\rho^4 \\ \alpha = -\rho \\ \beta = 2\rho \\ \gamma = 3\rho \end{array}\right\} \quad \text{luego} \quad \begin{array}{l} \alpha = 1/5 \\ \beta = -2/5 \\ \gamma = -3/5 \end{array}$$

Así, pues, la ecuación reducida que se pedía era:

$$\frac{x^2}{5} - \frac{y^2}{5/2} - \frac{z^2}{5/3} = 1 \quad \text{o} \quad x^2 - 2y^2 - 3z^2 = 5$$

y la cuádrica es, entonces, un hiperboloide de dos hojas (no reglado).

ECUACION REDUCIDA DE UNA CUADRICA CON CENTRO(*)

[342]

En el espacio afín euclídeo tridimensional y respecto de una referencia rectangular, se considera la cuádrica con centro cuya matriz es M; sea $A = [a_{ij}]$ su submatriz de los términos cuadráticos (nótese que $\det A \neq 0$, pues la cuádrica tiene centro). Se llaman *ejes de la cuádrica* a las tres rectas (perpendiculares entre sí) que pasan por su centro y tienen las direcciones de los vectores de una base ortonormal de vectores propios de la matriz A.

La ecuación de la cuádrica (con centro) respecto de sus ejes (referencia rectangular; coordenadas x, y, z) se llama *ecuación reducida* de la cuádrica y es la siguiente:

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + \frac{\det M}{\det A} = 0 \quad \begin{pmatrix} \lambda_1, \lambda_2 \text{ y } \lambda_3 \text{ son los} \\ \text{autovalores de } A \end{pmatrix} \qquad [1]$$

Los ejes de una cuádrica (con centro) son ejes de simetría de la misma.

DEMOSTRACION

Vale aquí lo dicho en la demostración de [318], con los retoques pertinentes, salvo el último párrafo, en el que se habla de «pendientes» (m) de rectas, que no es aplicable a nuestro caso (dimensión 3). Donde allí se habla de: cónica, vectores propios $\bar{u}_1$ y $\bar{u}_2$, autovalores λ_1 y λ_2, i y j variando de 1 a 2 o $\lambda_1 x^2 + \lambda_2 y^2$, habrá que poner aquí: cuádrica, vectores propios $\bar{u}_1$, $\bar{u}_2$ y $\bar{u}_3$, autovalores λ_1, λ_2 y λ_3, i y j variando de 1 a 3 o $\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2$. Las referencias [313], [316] y [317] deben ser sustituidas por las [335], [340] y [341].

EJERCICIO

Hallar la ecuación reducida de la cuádrica

$$x^2 + y^2 + z^2 + 6(xy + xz + yz) = 2$$

RESOLUCION

Para esta cónica es (con la notación usual, la de [342]):

$$\det M = -56 \quad , \quad \det A = 28 \quad , \quad \lambda_1 = \lambda_2 = 2 \quad , \quad \lambda_3 = 7$$

(*) Elipsoides, hiperboloides y conos.

Se trata, pues, de un elipsoide, cuya ecuación reducida será

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + \frac{\det M}{\det A} = 0 \quad \text{o sea} \quad 2x'^2 + 2y'^2 + 7z'^2 = 2$$

ECUACION REDUCIDA DE UN PARABOLOIDE

[343]

En el espacio afín euclídeo tridimensional y respecto de una referencia rectangular, se considera el paraboloide cuya matriz es M; sea $A = [a_{ij}]$ su submatriz de los términos cuadráticos, uno de cuyos autovalores es $\lambda = 0$ (pues $\det A = 0$, dado que hablamos de un paraboloide), y sean λ_1 y λ_2 sus otros autovalores. Llamemos $\bar{\delta}$ a un vector de la dirección del paraboloide (*), que es vector propio de A correspondiente al autovalor $\lambda = 0$; sean $\bar{u}_1$ y $\bar{u}_2$ vectores propios correspondientes a λ_1 y λ_2 (los $\bar{\delta}$, $\bar{u}_1$ y $\bar{u}_2$ forman base ortogonal).

En el paraboloide hay un punto V, llamado *vértice*, en el que el plano tangente es perpendicular a $\bar{\delta}$. Se llama *eje* (**) y *seudoejes* del paraboloide a las rectas que pasan por V y tienen las direcciones de $\bar{\delta}$, $\bar{u}_1$ y $\bar{u}_2$.

La ecuación del paraboloide respecto de sus seudoejes y de su eje (referencia rectangular; coordenadas x, y, z) se llama ecuación reducida y es:

$$\lambda_1 x^2 + \lambda_2 y^2 = 2pz \quad \text{con} \quad p = \sqrt{-\frac{\det M}{\lambda_1 \lambda_2}}$$

(*) Según ya se sabe (véase [337], si $\delta_0 = 0$, δ_1, δ_2, δ_3 son los adjuntos de los elementos de la primera fila de M, entonces $\bar{\delta}(\delta_1, \delta_2, \delta_3)$.

(**) El eje es la recta de intersección de los planos polares de dos puntos del infinito de direcciones perpendiculares a $\bar{\delta}$ (por ejemplo, las de $\bar{u}_1$ y $\bar{u}_2$), ya que ambos planos polares pasan por el vértice y por el punto de tangencia del paraboloide con el plano del infinito.

DEMOSTRACION

Sean P_1 y P_2 los puntos del infinito de las direcciones de los vectores $\bar{u}_1$ y $\bar{u}_2$; si U_1 y U_2 son las columnas de coordenadas de $\bar{u}_1$ y $\bar{u}_2$, entonces las columnas de coordenadas homogéneas de los puntos P_1 y P_2 son

$$\begin{bmatrix} 0 \\ U_1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 \\ U_2 \end{bmatrix}$$

Conviene no olvidar que, como $\bar{u}_1$ y $\bar{u}_2$ son vectores propios de A correspondientes a los autovalores λ_1 y λ_2, se verifica que $AU_1 = \lambda_1 U_1$ y que $AU_2 = \lambda U_2$.

Los planos polares (véase [329], III) de P_1 y P_2 son los planos cuyos coeficientes (de sus ecuaciones) forman las matrices fila(*):

$$[0\ U_1^t]M \quad \text{y} \quad [0\ U_2^t]M \qquad [1]$$

(*) En [329], III, estas matrices fila se denotan por $\alpha^t M$, lo que da lugar a la ecuación $\alpha^t M X = O$ de la que allí se habla.

Expresando la matriz M en bloques (en función de la matriz A, de la matriz B de los términos lineales y del término independiente c), se tiene que las matrices fila [1] se pueden poner (para $i = 1$ e $i = 2$) como sigue:

$$[0\ U_i^t]M = [0\ U_i^t]\begin{bmatrix} c & B^t \\ B & A \end{bmatrix} = [U_i^tB\ U_i^tA] = [U_i^tB\ \lambda U_i^t]$$

Luego $\lambda_1 U_1^t$ y $\lambda_2 U_2^t$ son las filas de coordenadas de los vectores perpendiculares a los planos polares de P_1 y P_2; por tanto, estos planos son perpendiculares a los vectores $\lambda_1\bar{u}_1$ y $\lambda_2\bar{u}_2$ y, por tanto se cortan según una recta, e, que tiene la dirección del vector $\bar{\delta}$, esto es, con la dirección del paraboloide. La recta e, como corta al paraboloide en su punto del infinito, le corta en otro punto, V. Veamos que V es el vértice y, por tanto, e es el eje. El punto V pertenece a los dos cilindros circunscrito al paraboloide desde P_1 y P_2, luego las rectas que pasan por V y tienen las direcciones de $\bar{u}_1$ y $\bar{u}_2$ son tangentes al paraboloide y, por ello, determinar el plano tangente en V que es, pues, perpendicular a $\bar{\delta}$, lo que confirma que V es el vértice.

Nos encontramos pues con que los seudoejes y el eje constituyen ejes cartesianos rectangulares y, además, del tipo que se considera en el Ejercicio [338], luego en dichos ejes la ecuación del paraboloide es del tipo

$$\alpha x^2 + 2\beta xy + \gamma y^2 = 2pz$$

(podría dividirse toda la ecuación por p, como se hizo en [338], pero aquí no es conveniente). Como, en esta referencia. los vectores (1, 0, 0) y (0, 1, 0) son vectores propios, correspondientes a λ_1 y λ_2, de la nueva matriz A, se deduce de ello fácilmente que $\gamma = 0$ y se pueden tomar $\alpha = \lambda_1$ y $\beta = \lambda_2$ (o proporcionales a ellos), con lo que la ecuación queda

$$\lambda_1 x^2 + \lambda_2 y^2 = 2pz$$

Recurriendo, finalmente, a que $\det M$ es invariante, se concluye que:

$$\det M = -p^2\lambda_1\lambda_2, \quad \text{luego} \quad p^2 = -\frac{\det M}{\lambda_1\lambda_2}$$

EJERCICIO

Hallar la ecuación reducida de la siguiente cuádrica, en función del valor del parámetro $\alpha \in \mathbb{R}$:

$$2x_1^2 + x_2^2 + 2x_3^2 + 2\alpha^2 x_1x_3 + 4x_1 + 2 = 0$$

RESOLUCION

Vamos a recurrir a los invariantes métricos, para lo que necesitamos de la matriz M de la cuádrica y de la submatriz A de los términos cuadráticos, así como de los autovalores λ_1, λ_2 y λ_3 de esta, que son:

$$M = \left[\begin{array}{c|ccc} 2 & 2 & 2 & 0 \\ \hline 2 & \alpha^2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \alpha^2 & 0 & 2 \end{array}\right] = \left[\begin{array}{c|c} 2 & *** \\ \hline * & \\ * & A \\ * & \end{array}\right] \quad \begin{array}{l} \det M = -2\alpha^4; \quad \det A = 4 - \alpha^4 \\ \lambda_1 = 1, \quad \lambda_2 = 2 + \alpha^2, \quad \lambda_3 = 2 - \alpha^2 \end{array}$$

Por tanto, según los valores de α^2, se tiene:

1. Si $\alpha^2 \neq 2$, es $\det A \neq 0$ y, por ello, la cuádrica tiene centro y su ecuación reducida es (véase [342]):

$$x^2 + (2 + \alpha^2)y^2 + (2 - \alpha^2)z^2 = \frac{2\alpha^4}{4 - \alpha^4}$$

   Para $\alpha^2 = 0$, la cuádrica $(x^2 + 2y^2 + 2z^2 = 0)$ sólo tiene un punto, el origen; se dice que es un «cono imaginario» con vértice real. Si $0 < \alpha^2 < 2$, la cuádrica es un elipsoide. Si $2 < \alpha^2$, la cuádrica es un hiperboloide de dos hojas.

2. Si $\alpha^2 = 2$, hay un autovalor nulo $(\lambda_3 = 0)$ y la cuádrica es no degenerada $(\det M = -8 \neq 0)$, luego es un paraboloide y su ecuación reducida es (véase [343]):

$$x^2 + 4y^2 = 2\sqrt{2}\,z$$

   Esta cuádrica es, pues, un paraboloide elíptico.

Ejercicios y problemas a la parte VI

ENUNCIADOS

VI.1. Se considera la elipse de semiejes a y b $(a>b)$ y cuyos vértices del semieje menor son los puntos B y B'. Si X es un punto de esta elipse y P y Q son los puntos en los que las rectas XB y XB' cortan al eje mayor de la elipse, pruébese que $K=\overline{OP}\cdot\overline{OQ}$ es constante (O es el centro de la elipse), esto es, que no depende de X.

VI.2. Sea dada una elipse y sea F uno de sus focos. Hallar el lugar geométrico que describe el punto, P, mitad del segmento FX, al hacer que X recorra la elipse.

VI.3. Considérese una hipérbola, de ecuación reducida $x^2/a^2-y^2/b^2=1$:

1. Hallar la distancia de un foco F de la hipérbola a una de sus asíntotas.
2. Compruébese que una asíntota, la perpendicular a dicha asíntota desde un foco y la correspondiente directriz pasan por una mismo punto.

VI.4. Considérese una hipérbola de semiejes a (real) y b (imaginario). Sean d_1 y d_2 las distancias de un punto X a las asíntotas de la hipérbola. Pruébese que $k=d_1d_2$ es constante para cualquiera que sea el punto X de la hipérbola; hallar k.

VI.5. Se considera una elipse de semiejes a (mayor) y b (menor); sean A y A' sus vértices del eje mayor, sea F uno de sus focos y sea d la correspondiente directriz. Sea X un punto cualquiera de la elipse y sean P y Q los puntos en los que las rectas XA y XA' cortan a la directriz d. Compruébese que FP y FQ son perpendiculares.

VI.6. Sean dados un punto A y una recta r que no pasa por A. Si X es un punto, del plano de A y r, sea X' su proyección ortogonal sobre r. Hallar el lugar geométrico que describe el punto X cuando se mueve de manera que $c\overline{XX'}=\overline{AX'}$, donde $c>0$ es un número dado.

VI.7. Se considera una elipse de semiejes a (mayor) y b (menor). Sea XX' una cuerda de la elipse que se mueve manteniéndose paralela al eje menor y considérense las rectas XA y $X'A'$, donde A y A' son los vértices situados en el eje mayor. Hallar el lugar geométrico que describe el punto P de intersección de XA y $X'A'$.

VI.8. Se considera la parábola que tiene por ecuación reducida a $y^2=2px$. Sean X y X' dos puntos cualesquiera de la parábola tales que el ángulo XOX' es recto (O es el origen). Pruébese que XX' corta al eje de la parábola en un punto P fijo.

VI.9. Se considera una parábola de parámetro p. Hallar el lugar geométrico descrito por los puntos medios de las cuerdas XX' de la parábola que pasan por su foco F.

VI.10. Se considera la parábola de ecuación reducida $y^2=2px$ y se traza la normal (perpendicular a la tangente) en un punto X de ella; sea Y el punto en el que esta normal corta al eje de la parábola. Hallar el lugar geométrico del punto X' simétrico de X respecto de Y, cuando X recorre la parábola.

VI.11. Considérese la parábola que tiene su foco en F y tiene por directriz a la recta d. Comprobar que la tangente a la parábola en un punto X, cualquiera, de ella es la bisectriz de $\overrightarrow{XF}$ y $\overrightarrow{XX'}$, donde X' es la proyección de X sobre d.

VI.12. Considérese una parábola y un punto X exterior a ella; por X se trazan las tangentes a la parábola, que son tangentes en los puntos P_1 y P_2; sean P, Q_1 y Q_2 los puntos medios de P_1P_2, XP_1 y XP_2. Pruébese que:

1. XP es paralela al eje de la parábola.
2. Q_1Q_2 es tangente a la parábola y lo es en su punto medio Q.

VI.13. Una hebra de hilo anudada en sus extremos y de longitud $2p$, se mantiene formando triángulo isósceles OXX', donde O es fijo y XX' permanece paralelo a cierta dirección fija. Hallar los lugares geométricos que describen los puntos X y X'.

VI.14. Se da una circunferencia fija, con centro en O y radio r; también se considera un punto fijo, P,

siendo $\overline{OP} = 2a$. Hallar el lugar geométrico descrito por los centros, X, de las circunferencias tangentes a la circunferencia dada y que pasan por P (indicación: tómese el origen de coordenadas en el punto medio de OP).

VI.15. Dos puntos X_1 y X_2 se mueven, con la misma velocidad angular, en sendas circunferencias concéntricas de radios r_1 y r_2. En el instante inicial los puntos móviles están alineados con el centro y a un mismo lado de éste; se mueven en sentidos contrarios. Hallar la ecuación del lugar geométrico que describe el punto medio, X, de los X_1 y X_2.

VI.16. Compruébese que la tangente a una elipse en uno de sus puntos X es la bisectriz exterior del ángulo FXF' (F y F' son los focos de la elipse).

VI.17. Dada una elipse y un punto P de ella, determinar el triángulo de área máxima de entre los que tienen un vértice en P y los otros dos también están en la elipse.

VI.18. Si e_1 y e_2 son las excentricidades de dos hipérbolas conjugadas, compruébese que $(1/e_1^2) + (1/e_2^2) = 1$.

VI.19. Pruébese que si los tres vértices de un triángulo están situados en una hipérbola equilátera, entonces el ortocentro del triángulo también está en la hipérbola.

VI.20. En un punto A se efectúa un disparo; el proyectil va a velocidad v y hace impacto en un punto B. Hallar el lugar geométrico de los puntos X desde los que se escucha simultáneamente el disparo y el impacto (llámese v' a la velocidad del sonido).

VI.21. Se consideran dos hipérbolas conjugadas, de semiejes a y b. Se traza una tangente común a ambas. Hallar la distancia entre los puntos de tangencia, comprobando que es constante.

VI.22. Hallar los planos tangentes a una cuádrica que son paralelos a un plano dado $ax + by + cz = 1$ (considérese una cuádrica con centro $\alpha x^2 + \beta y^2 + \gamma z^2 = 1$ y también un paraboloide $2z = \alpha x^2 + \beta y^2$).

VI.23. Indicar cómo se determinan los planos tangentes a una cuádrica que contienen a una recta dada. Supóngase que la recta y la cuádrica son:

$$r: \begin{cases} ax + by + cz = 1 \\ a'x + b'y + c'z = 1 \end{cases} \qquad C: \alpha x^2 + \beta y^2 + \gamma z^2 = 1$$

VI.24. Recurriendo a las identidades:

$$(1 + \alpha^2 + \beta^2)^2 = (1 - \alpha^2 - \beta^2)^2 + (2\alpha)^2 + (2\beta)^2$$

$$(1 - \alpha^2 - \beta^2) = (1 + \alpha^2 + \beta^2)^2 - (2\alpha)^2 - (2\beta)^2$$

$$(\alpha + \beta)^2 = (1 + \alpha\beta)^2 + (\alpha - \beta)^2 - (1 - \alpha\beta)^2$$

Comprobar que el elipsoide, el hiperboloide de dos hojas y el hiperboloide de una hoja admiten, respectivamente, las siguientes ecuaciones paramétricas:

1. $$\frac{x}{a} = \frac{1 - \alpha^2 - \beta^2}{1 + \alpha^2 + \beta^2} \quad , \quad \frac{y}{b} = \frac{2\alpha}{1 + \alpha^2 + \beta^2}$$

$$\frac{z}{c} = \frac{2\beta}{1 + \alpha^2 + \beta^2} \quad \text{(elipsoide)}$$

2. $$\frac{x}{a} = \frac{1 + \alpha^2 + \beta^2}{1 - \alpha^2 - \beta^2} \quad , \quad \frac{y}{b} = \frac{2\alpha}{1 - \alpha^2 - \beta^2}$$

$$\frac{z}{c} = \frac{2\beta}{1 - \alpha^2 - \beta^2} \quad \text{(hiperboloide de dos hojas)}$$

3. $$\frac{x}{a} = \frac{1 + \alpha\beta}{\alpha + \beta} \quad , \quad \frac{y}{b} = \frac{\alpha - \beta}{\alpha + \beta}$$

$$\frac{z}{c} = \frac{1 - \alpha\beta}{\alpha + \beta} \quad \text{(hiperboloide de una hoja)}$$

VI.25. Demuéstrese, en una cuádrica con centro, que la suma de los cuadrados de las distancias del centro a tres planos tangentes ortogonales es constante; hállese dicha constante.

VI.26. Considérese un paraboloide y tres planos tangentes a él que son ortogonales entre sí. Sean P_1, P_2 y P_3 las proyecciones del vértice O sobre dichos tres planos. Pruébese que la suma de las proyecciones sobre el eje de $\overline{OP_1}$, $\overline{OP_2}$ y $\overline{OP_3}$ es constante; hállese esta constante.

VI.27. Dada una cuádrica con centros $\alpha x^2 + \beta y^2 + \gamma z^2 = 1$, pruébese que el lugar geométrico que forman los puntos desde los que es posible trazar tres planos tangentes a la cuádrica que sean ortogonales dos a dos (forman triedro trirrectángulo circunscrito a la cuádrica), es la esfera $x^2 + y^2 + z^2 = 1/\alpha + 1/\beta + 1/\gamma$ (esfera de Monge).

VI.28. Dado un paraboloide $2z = \alpha x^2 + \beta y^2$, pruébese que el lugar geométrico que forman los puntos

desde los que es posible trazar tres planos tangentes al paraboloide que sean ortogonales dos a dos, es al plano $z=(\alpha+\beta):(2\alpha\beta)$, perpendicular al eje del paraboloide.

VI.29. Considérese una cuádrica con centro y sean d_1, d_2 y d_3 las longitudes de tres de sus semidiámetros (los semidiámetros son los segmentos que unen el origen con los puntos de la cuádrica). Pruébese que si los tres semidiámetros son ortogonales dos a dos, entonces $1/d_1^2+1/d_2^2+1/d_3^2$ es constante.

VI.30. Se consideran cuatro puntos alineados A, B, C y P; sean $a=\overline{PA}$, $b=\overline{PB}$ y $c=\overline{PC}$ (fijos). El segmento $ABCP$ se mueve de manera que A, B y C se desplazan, respectivamente, por los planos $x=0$, $y=0$ y $z=0$. Hállese el lugar geométrico que describe el punto P.

VI.31. Dada la hipérbola $(x^2/a^2)-(y^2/b^2)=1$, $z=0$, hallar el lugar geométrico de los vértices X de los triedros trirrectangulares cuyas aristas cortan a la hipérbola.

VI.32. Hallar el lugar geométrico de los puntos X tales que su distancia al origen y su distancia al plano $z=d$ están en la relación e (constante). Discutir el resultado en función de los valores de $e>0$.

VI.33. Se considera una esfera de radio r; sean A y B dos puntos diametralmente opuestos de la esfera; sean t_a y t_b rectas tangentes a la esfera en A y B, que son perpendiculares entre sí. Hallar el lugar geométrico descrito por las rectas tangentes a la esfera y que cortan a t_a y t_b. Hallar la intersección de esta superficie y la esfera.

VI.34. Considérese una cualquiera de las generatrices rectilíneas de un hiperboloide de una hoja. Probar que la proyección de la generatriz sobre el plano de la elipse de garganta es tangente a dicha elipse.

VI.35. Considérese una cualquiera de las generatrices rectilíneas del paraboloide hiperbólico $2z=(x^2/p)-(y^2/q)$. Probar que la proyección de la generatriz sobre el plano $y=0$ es tangente a la parábola del paraboloide situada en dicho plano.

VI.36. Se considera el paraboloide hiperbólico (reglado) $2z=x^2/p-y^2/q$ y sus dos planos principales ($x=0$ e $y=0$). Una generatriz cualquiera del hiperboloide corta a sus planos principales en los puntos P y Q; hallar el lugar geométrico que describe el punto X medio de los P y Q.

VI.37. Clasificar las siguientes cónicas, del plano afín:

a) $2x^2+y^2+2xy-12x-4y+3=0$.
b) $x^2+3y^2+4xy+4x-2y-4=0$.
c) $x^2+y^2-2xy+2y+1=0$.

VI.38. Clasificar las siguientes cónicas, del plano afín:

a) $2x^2+y^2-2xy+2x-2y+1=0$.
b) $4x^2+y^2+4xy-y=0$.
c) $2x^2-y^2+xy+3y-2=0$.
d) $x^2+y^2-2y+2=0$.

VI.39. Clasificar la siguiente cónica, del plano afín, en función de los parámetros a; $b\in\mathbb{R}$.

$$(1+a)x^2+2axy+ay^2+2bx+a-3b^2=0$$

VI.40. Clasificar y hallar la ecuación reducida de la siguiente cónica, del plano euclídeo:

$$11x_1^2-4x_1x_2+14x_2^2+40x_1+20x_2+45=0$$

Hallar también el centro y los ejes de dicha cónica.

VI.41. Clasificar y hallar la ecuación reducida de la siguiente cónica, del plano euclídeo:

$$4x_1^2+4x_1x_2+x_2^2+4x_1-8x_2-20=0$$

Hallar los ejes en los que la ecuación es la reducida.

VI.42. Hallar la ecuación reducida, en función de θ, de la cónica del plano euclídeo:

$$x_1^2\cos\theta-2x_1x_2+x_2^2\cos\theta+2x+2y=0$$

$(-\pi<\theta<\pi)$. Determinar los ejes en los que la ecuación es la reducida. Clasificar la cónica.

VI.43. Se considera una parábola, cuyo parámetro p (distancia del foco a la directriz) es fijo, que gira manteniendo su foco inmóvil. Se considera también una cierta dirección fija en el plano de la parábola y se traza la tangente a ésta que tiene la dirección dada. Hallar la ecuación, en ejes *ad hoc*, del lugar geométrico que describe el punto de tangencia.

VI.44. Se considera una cónica ordinaria con centro; sea M su matriz y sea $A=[a_{ij}]$ la submatriz de térmi-

tancia constannos cuadráticos. Compruébese que los coeficientesα y β de la ecuación reducida, $\alpha x^2 + \beta y^2 = 1$, de la cónica son las raíces de la ecuación de segundo grado:

$$(\det M)^2\lambda^2 + (\det M)\,(\det A)(a_{11} + a_{22})\lambda + \\ + (\det A)^3 = 0$$

VI.45. Comprobar que, si F es foco de una cónica C, las tangentes desde F a C son imaginarias y tienen pendientes $+i$ y $-i$ (se llaman las rectas isótropas que pasan por F).

VI.46. Hallar la ecuación del hiperboloide que tiene el mismo cono asintótico que:

$$C: x^2 - y^2 + 4xz - 2yz + 10z + 1 = 0$$

y que es tangente al plano π: $2x - y + z = 0$.

VI.47. Considérese la cuádrica que, en cierta referencia cartesiana del plano afín, admite por ecuación a:

$$x^2 - y^2 + pz^2 + 2x + 2(p-1) + p = 0$$

($p \in \mathbb{R}$ parámetro).

1. Hallar p para que la cuádrica sea un par de planos y hallar éstos.
2. Hallar los valores de p para los que la cuádrica es reglada.
3. Hallar, si la cuádrica es reglada, sus dos generatrices rectilíneas que pasan por el punto $A(-1, -1, 1)$ de la cuádrica.

VI.48. Se considera la siguiente cuádrica C del espacio afín:

$$C: 2x^2 - y^2 - 3z^2 - 2xz - 4y + 10 = 0$$

1. Hallar el lugar geométrico que engendran las tangentes a C en el punto $P_1(1, 2, 0)$ de ella.
2. Hallar el lugar geométrico que engendran las tangentes a C desde el punto $P_2(0, 2, 0)$, que no es de C. Hallar el plano en el que se encuentran los puntos de contacto con C de dichas tangentes.
3. Comprobar que el plano $z = 2$ es tangente a C y hallar su punto de tangencia.
4. Analizar si existe algún plano paralelo al $x = 0$ que sea tangente a la cuádrica C.

VI.49. Sea C una cuádrica ordinaria que, en cierta referencia cartesiana tiene a M por matriz. Se llama *polo* de un plano al punto cuyo plano polar es el plano dado. Hallar la columna de coordenadas del polo de un plano en función de la columna U de los coeficientes u_0, u_1, u_2, u_3 de la ecuación del plano dado. Hallar la condición, a cumplir por U, que caracteriza a los planos tangentes a la cuádrica.

VI.50. Hallar la ecuación del lugar geométrico que describe el centro de la siguiente cuádrica, al variar los parámetros a y b:

$$x^2 + y^2 - z^2 + axz + byz - 2x - 8y + 4z = 0$$

VI.51. Clasificar las siguientes cuádricas, del espacio afín:

a) $x^2 + y^2 + 4z^2 + 2xy + 12xz - 4yz + 2x - 6y - \\ - 4z + 1 = 0$.
b) $x^2 + 2y^2 + z^2 + 2xz + 2x + 1 = 0$.
c) $-x^2 + y^2 - 2z^2 - 6xz + 2yz - 2x - 6y = 2$.

VI.52. Clasificar las siguientes cuádricas, del espacio afín:

a) $y^2 + z^2 - 2yz + 6x = 4$.
b) $5x^2 + 4y^2 - 2z^2 + 6xy + 4xz + 6x + 8y + 3 = \\ = 0$.
c) $3x^2 + 2y^2 + 6xy + 8x + 4y + 2z + 2 = 0$.

VI.53. Clasificar la siguiente cuádrica, del espacio afín, en función del valor que tome el parámetro α:

$$2x^2 + y^2 + 2z^2 + 2xy + 2\alpha yz - 4y + 1 = 0$$

VI.54. Clasificar, en función del parámetro p, la siguiente cuádrica, del espacio afín:

$$x^2 + py^2 + (p-1)z^2 + 2xy - 2yz + \\ + 2x + 2z + 4 = 0$$

VI.55. Clasificar, en función de los parámetros a y b, la siguiente cuádrica, del espacio afín

$$3x^2 + 2y^2 + az^2 + 4xy + 2x + 2bz + 1 = 0$$

VI.56. Hallar la ecuación reducida de la siguiente cuádrica, del espacio euclídeo, y clasificarla:

$$6x_1^2 + 3x_2^2 + 3x_3^2 + 4x_1x_2 - 4x_1x_3 - \\ - 2x_2x_3 + 4x_1 - 8x_2 + 9 = 0$$

Hallar su centro y sus ejes. Esta cuádrica, ¿es de revolución?

VI.57. Considérense las cuádricas que, en ejes rectangulares, admiten ecuación del tipo

$$\alpha x_1^2 + (1-\alpha)x_2^2 + \alpha x_3^2 + 2(1-\alpha)x_1x_3 + \\ + 2x_1 + 2x_3 + 3 = 0 \quad (\alpha \in \mathbb{R})$$

Clasificar las cuádricas dadas, en función de $\alpha \in \mathbb{R}$. Hallar los valores de α para los que se obtiene una cuádrica de revolución. Hallar la ecuación reducida de la cuádrica correspondiente a $\alpha = 3$; determinar sus ejes.

VI.58. Sea $C(p)$ la cuádrica, dependiente del parámetro p, que en cierta referencia cartesiana rectangular tiene por ecuación a:

$$4x_1^2 + 2(p-1)x_2^2 + 4x_3^2 + 4x_1x_3 + 2px_2 + p = 0$$

a) Clasificar la cuádrica $C(P)$, en función de p.

b) Obtener la ecuación reducida de la cuádrica $C(-1)$.

c) Obtener la ecuación reducida de aquella de las cuádricas $C(p)$ que es paraboloide.

d) Obtener la referencia correspondiente a la anterior ecuación.

VI.59. *a*) Hallar (en ejes rectangulares xyz) la ecuación general de todos los paraboloides que contienen a la parábola $y^2 = 2x$, $z = 0$. *b*) De entre ellos, determinar los que son de revolución. *c*) Hallar el lugar geométrico descrito por los vértices de estos últimos.

SOLUCIONES

VI.1. $X(\alpha, \beta)$ con $\alpha^2/a^2 + \beta^2/b^2 = 1$; $\overline{OP} = b\alpha : (b - \beta)$ y $\overline{OQ} = b\alpha : (b + \beta)$; $\overline{OP} \cdot \overline{OQ} = (b^2\alpha^2) : (b^2 - \beta^2) = a^2$.

VI.2. Elipse de semiejes a y b; distancia focal $2c$. Si $P(x, y)$, entonces $X(2x - c, 2y)$ satisface a la ecuación de la elipse, luego

$$\frac{(x - c/2)^2}{(a/2)^2} + \frac{y^2}{(b/2)^2} = 1$$

El lugar en la elipse de semiejes mitad de los de la elipse E dada, que tiene por centro al punto medio de F y del centro de E y con igual eje mayor que E.

VI.3. 1. Foco $(c, 0)$, asíntota $bx - ay = 0$, distancia $|-cb| : \sqrt{a^2 + b^2} = b$.

2. Asíntota $bx - ay = 0$, perpendicular $ax + by = ac$, directriz $x = a^2/c$; las tres pasan por $(a^2/c, ba/c)$.

VI.4. $X(\alpha, \beta)$ con $\alpha^2/a^2 - \beta^2/b^2 = 1$

$$k = d_1 \cdot d_2 = \frac{|\alpha/a - \beta/b|}{\sqrt{1/a^2 + 1/b^2}} \cdot \frac{|\alpha/a + \beta/b|}{\sqrt{1/a^2 + 1/b^2}} =$$

$$= \frac{1}{1/a^2 + 1/b^2} = \frac{a^2b^2}{a^2 + b^2}$$

VI.5. $X(\alpha, \beta)$ con $\alpha^2/a^2 + \beta^2/b^2 = 1$; $F(c, 0)$ con $c^2 = a^2 + b^2$; $d : x = a^2/c$

$$P\left(\frac{a^2}{c}, \frac{a(a - c)\beta}{c(\alpha - a)}\right) \quad , \quad Q\left(\frac{a^2}{c}, \frac{a(a + c)\beta}{c(\alpha + a)}\right)$$

$$\overline{FP} \cdot \overline{FQ} = \frac{(a^2 - c^2)^2}{c^2} + \frac{a^2(a^2 - c^2)\beta^2}{c^2(\alpha^2 - a^2)} =$$

$$= \frac{b^2}{c^2}\left(b^2 + \frac{a^2b^2 - b^2\alpha^2}{\alpha^2 - a^2}\right) = 0$$

VI.6. Tomando a r como «eje de la y» y el «eje de la x» pasando por A, con lo que $A(a, 0)$. Lugar $c|x| = \sqrt{a^2 + y^2}$, o sea $c^2x^2 - y^2 = a^2$, hipérbola de semiejes a/c y a.

VI.7. $X(\alpha, \beta)$, $X' = (\alpha, -\beta)$ con $\alpha^2/a^2 + y^2/b^2 = 1$.

$$AX : \beta x + (a - \alpha)y = a\beta$$

$$A'X' = \beta x + (a + \alpha)y = -a\beta$$

Eliminando α y β entre las tres ecuaciones,

$$x^2/a^2 - y^2/b^2 = 1$$

VI.8. $X(\alpha^2/2p, \alpha)$, $X'(\beta^2/2p, \beta)$

$$\overrightarrow{OX} \cdot \overrightarrow{OX'} = 0 \quad , \quad \alpha\beta = -4p^2 \quad ; \quad P(h, 0)$$

con

$$h = \frac{\alpha^2}{2p} + \frac{\alpha}{\alpha - \beta}\,\frac{\beta^2 - \alpha^2}{2p} = \frac{-\alpha\beta}{2p} = 2p$$

El punto $P(2p, 0)$ es fijo.

VI.9. $y^2 = 2px$, $F(p/2, 0)$, $X(\alpha^2/2p, \alpha)$, $X'(\beta^2/2p, \beta)$. X, X', F alineados, luego

$$\left(\frac{\alpha^2}{2p} - \frac{p}{2}\right) : \alpha = \left(\frac{\beta^2 - \alpha^2}{2p}\right) : (\beta - \alpha)$$

o sea, $\alpha\beta = -p^2$. Punto medio:

$$x = \frac{\alpha^2 + \beta^2}{4p} \quad , \quad y = \frac{\alpha + \beta}{2}$$

Eliminando α y β entre las tres últimas ecuaciones, se obtiene $y^2 = px - p^2/2$ (parábola de parámetro mitad que la dada, con el mismo eje y con vértice en el foco de aquélla).

VI.10. $X(x_0, y_0)$ con $y_0^2 = 2px_0$; la normal es $y_0x + py = y_0x_0 + py_0$; $Y(x_0 + p, 0)$; $X'(x_0 + 2p, y_0)$. Lugar $y^2 = 2p(x - 2p)$.

VI.11. Parábola $y^2 = 2px$, $F(p/2, 0)$, $X(x_0, y_0)$, $X'(-p/2, y_0)$. Como $XF = XX'$, la bisectriz es la mediatriz de $X'F$, que es $-px + y_0(y - y_0/2) = 0$, que coincide con la tangente $yy_0 = px + px_0$.

VI.12. Parábola $y^2 = 2px$, $X(\alpha, \beta)$:

1. P_1, P_2: $y^2 = 2px$, $y\beta = px + p\alpha$;

$$P\left(\frac{\beta^2}{p} - \alpha, \beta\right)$$

luego $\overrightarrow{XP}$ paralelo al eje.

2. Q punto medio de XP,

$$Q\left(\frac{\beta^2}{2p}, \beta\right)$$

pertenece a la parábola y la tangente en él es $y\beta = px + p\beta^2/2p$, que es paralela a P_1P_2, luego es Q_1Q_2.

VI.13. Tómense ejes rectangulares Oxy con Ox paralelo a la dirección fija.

$$\overline{OX} + \overline{XX'} + \overline{X'O} = 2p \ ; \ X(x, y) \ , \ X'(-x, y);$$

$2\sqrt{x^2+y^2} + 2x = 2p$; $y^2 = -2px + p^2$ lugar de X; $y^2 = 2px + p^2$ lugar de X'; ambos para $-p \leqslant y \leqslant p$.

VI.14. OP = «eje de la x» con origen en el punto medio de OP; $O(-a, 0)$, $P(a, 0)$, $X(x, y)$;

$$\sqrt{(x-a)^2+y^2} = r - \sqrt{(x+a)^2+y^2};$$

$4(r^2 - 4a^2)x^2 + 4r^2y^2 = r^4 - 4r^2a^2$, que es elipse o hipérbola según que P sea interior o exterior a la circunferencia; sus focos son O y P.

VI.15. $X_1(r_1 \cos \omega t, r_1 \operatorname{sen} \omega t)$, $X_2(r_2 \cos \omega t, -r_2 \operatorname{sen} \omega t)$, ω = velocidad angular.

$$X\begin{cases} x = \frac{1}{2}(r_1 + r_2) \cos \omega t \\ y = \frac{1}{2}(r_1 - r_2) \operatorname{sen} \omega t \end{cases}$$

$$X: \frac{4x^2}{(r_1+r_2)^2} + \frac{4y^2}{(r_1-r_2)^2} = 1$$

VI.16. $x^2/a^2 + y^2/b^2 = 1$, $X(\alpha, \beta)$, tangente

$$(\alpha x)/a^2 + (\beta y)/b^2 = 1$$

$$\overrightarrow{FX}(\alpha - c, \beta) \ , \ \overrightarrow{F'X}(\alpha + c, \beta)$$

$$\|\overrightarrow{FX}\| = (a^2 - \alpha c) : a \ , \ \|\overrightarrow{F'X}\| = (a^2 + \alpha c) : a$$

bisectriz interior

$$\frac{\overrightarrow{FX}}{\|\overrightarrow{FX}\|} + \frac{\overrightarrow{F'X}}{\|\overrightarrow{F'X}\|} = \frac{2a}{(a^4 - \alpha^2 c^2)}(\alpha b^2, \beta b^2)$$

es perpendicular a la tangente.

VI.17. Considérese la elipse como proyección de una circunferencia sobre un plano. En la circunferencia, el mayor área se obtiene si el triángulo es equilátero; al proyectar, se obtiene que el triángulo pedido es el siguiente: si O es el centro de la elipse y P' el punto diametralmente opuesto de P, el lado del triángulo opuesto a P es paralelo a la tangente en P trazada por el punto medio de P' y O.

VI.18. $e_1 = c/a$, $e_2 = c/b$, $c^2 = a^2 + b^2$, $c^2 = (c/e_1)^2 + (c/e_2)^2$.

VI.19. Hipérbola $xy = k$, $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ y $P_3(x_3, y_3)$. Altura correspondiente a P_3:

$$x_1x_2x_3(x - x_3) - kx_3y + k^2 = 0$$

que corta a la hipérbola (además de en P_3) en $P(x_0, y_0)$ con $x_0 = -k^2/(x_1x_2x_3)$. Las otras dos alturas también pasan por P (por simetría), que es, entonces el ortocentro.

VI.20. $(\overline{AX} - \overline{BX}) : r' = \overline{AB} : v$. Hipérbola con focos en A y B y constante $2a = \overline{AB}(v'/v)$.

VI.21.

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \ , \ \frac{y^2}{a^2} - \frac{x^2}{b^2} = 1$$

Tangentes a una y otra

$$\frac{xx_1}{a^2} - \frac{yy_1}{b^2} = 1 \quad \text{e} \quad \frac{yy_2}{a_2} - \frac{xx_2}{b^2} = 1$$

Como las tangentes son iguales y los puntos son de las hipérbolas, es:

$$\frac{x_1}{a^2} = -\frac{x_2}{b^2} \ , \ \frac{y_1}{b^2} = -\frac{y_2}{a^2}$$

$$\frac{x_1^2}{a^2} - \frac{y_1^2}{b^2} = 1 \ , \ \frac{y_2^2}{a^2} - \frac{x_2^2}{b^2} = 1$$

de donde

$$x_1 = \frac{-a^2}{\sqrt{a^2-b^2}} \ , \ y_1 = \frac{-b^2}{\sqrt{a^2-b^2}}$$

$$x_2 = \frac{b^2}{\sqrt{a^2-b^2}} \ , \ y_2 = \frac{a^2}{\sqrt{a^2-b^2}}$$

luego la distancia pedida vale

$$(a^2 + b^2)\sqrt{2} : \sqrt{a^2 - b^2}.$$

VI.22. Cuádrica con centro: $\alpha x x_0 + \beta y y_0 + \gamma z z_0 = 1$ ha de ser paralelo al plano dado,

$$\alpha x_0/a = \beta y_0/b = \gamma z_0/c \quad \text{y} \quad \alpha x_0^2 + \beta y_0^2 + \gamma z_0^2 = 1$$

luego

$$x_0 = \rho a/\alpha \quad , \quad y_0 = \rho b/\beta \quad \text{y} \quad z_0 = \rho c/\gamma$$

con

$$\rho = 1 : \sqrt{a^2/\alpha + b^2/\beta + c^2/\gamma}$$

y el plano tangente

$$ax + by + cz = \sqrt{a^2/\alpha + b^2/\beta + c^2/\gamma}$$

Igualmente, para el paraboloide

$$ax + by + cz + \frac{a^2\beta + b^2\alpha}{\alpha\beta c^2} = 0$$

VI.23. $\alpha x x_0 + \beta y y_0 + \gamma z z_0 = 1$ debe coincidir con

$$(a + \lambda a')x + (b + \lambda b')y + (c + \lambda c')z = 1 + \lambda$$

siendo $\alpha x_0^2 + \beta y_0^2 + \gamma z_0^2 = 1$;

$$\lambda^2\left(\frac{a'^2}{\alpha} + \frac{\beta'^2}{\beta} + \frac{c'^2}{\gamma} - 1\right) +$$

$$+ 2\lambda\left(\frac{aa'}{\alpha} + \frac{bb'}{\beta} + \frac{cc'}{\gamma} - 1\right) +$$

$$+ \left(\frac{a^2}{\alpha} + \frac{b^2}{\beta} + \frac{c^2}{\gamma} - 1\right) = 0$$

Para cada λ raíz de esta ecuación se obtiene uno de los planos tangentes.

VI.24. Dividiendo cada una de las identidades por sus primeros miembros, las paramétricas resultan evidentes.

VI.25. $\alpha x^2 + \beta y^2 + \gamma z^2 = 1$; plano tangente

$$\alpha x_0 x + \beta y_0 y + \gamma z_0 z = 1 \quad \text{o} \quad xu + yv + zw = h$$

(h distancia al origen, (u, v, w) unitario normal al plano); identificando, queda

$$h^2 = (u^2/\alpha) + (v^2/\beta) + (w^2/\gamma)$$

Si los tres planos tangentes son $u_i x + v_i y + w_i z = h_i$, sumando se obtiene $\Sigma h_i^2 = 1/\alpha + 1/\beta + 1/\gamma$.

VI.26. $2z = \alpha x^2 + \beta y^2$; plano tangente

$$\alpha x_0 x + \beta y_0 y - z - z_0 = 0 \quad \text{o} \quad xu + yv + zw = h$$

(h distancia al origen, (u, v, w) unitario normal al plano); identificando, queda

$$h = -(u^2/\alpha + v^2/\beta) : (2w)$$

la proyección de $\overline{OP}$ vale

$$hw = -(u^2/\alpha + v^2/\beta) : 2$$

Si los tres planos son $u_i x + v_i y + w_i z = h_i$, sumando las proyecciones. se obtiene $\Sigma w_i h_i = -(1/\alpha + 1/\beta) : 2$.

VI.27. Según se vio en el Ejercicio VI.25, tres planos tangentes perpendiculares entre sí son

$$xu_i + yv_i + zw_i = [u_i^2/\alpha + v_i^2/\beta + w_i^2/\gamma]^{1/2}$$

elevando al cuadrado y sumando, como los (u_i, v_i, w_i) forman base ortonormal, se obtiene la esfera de Monge.

VI.28. Según se vio en el Ejercicio VI.26, tres planos tangentes perpendiculares entre sí son

$$u_i x + v_i y + w_i z = -(u_i^2/\alpha + v_i^2/\beta) : (2w_i)$$

eliminando los parámetros (teniendo en cuenta que los (u_i, v_i, w_i) forman base ortonormal), se obtiene el plano del enunciado.

VI.29. $\alpha x^2 + \beta y^2 + \gamma z^2 = 1$. Si (u_i, v_i, w_i) es el unitario de la dirección del semidiámetro d_i, es

$$1/d_i^2 = \alpha u_i^2 + \beta v_i^2 + \gamma w_i^2$$

Teniendo en cuenta que los (u_i, v_i, w_i) forman base ortonormal, sumando se obtiene el resultado pedido.

VI.30. Si (u, v, w) es el unitario de la dirección del segmento, $P(au, bv, cw)$ como $u^2 + v^2 + z^2 = 1$, el lugar es el elipsoide $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$.

VI.31. $X(\alpha, \beta, \gamma)$; una arista $x = \alpha + \lambda u$, $y = \beta + \lambda v$, $z = \gamma + \lambda w$, (u, v, w) unitario; por cortar a la hipérbola

$$w^2(\alpha^2/a^2 - \beta^2/b^2 - 1) + u^2(\gamma^2/a^2) - v^2(\gamma^2/b^2) -$$

$$- 2uw(\alpha\gamma/a^2) + 2vw(\beta\gamma/b^2) = 0$$

Sustituyendo (u, v, w) por (u_i, v_i, w_i), donde estos vectores forman base ortonormal, eliminando los parámetros se obtiene

$$\alpha^2/a^2 - \beta^2/b^2 + z^2(1/a^2 - 1/b^2) = 1$$

VI.32. $x^2 + (1 - e^2)y^2 + 2de^2z = d^2e^2$. Si $e > 1$, es un hiperboloide de dos hojas; para $e = 1$, es un paraboloide elíptico; si es $e < 1$, es un elipsoide (todos ellos de revolución).

VI.33. Se toma AB como eje OZ, $A(0,0,r)$; t_a: $x = y$, $z = r$; t_b: $x = -y$, $z = -r$. La recta PQ con $P(\alpha, \alpha, r)$ y $Q(\beta, -\beta, -r)$ es tangente si $\alpha^2\beta^2 = r^4$. La superficie: $(x^2 - y^2)^2 = (z^2 - r^2)^2$, que son dos hiperboloides regladas: $x^2 - y^2 \pm z^2 = \pm r^2$. La intersección consta de dos circunferencias:

$$\begin{cases} x^2 + y^2 + z^2 = r^2 \\ x = 0 \end{cases} \quad \text{y} \quad \begin{cases} x^2 + y^2 + z^2 = r^2 \\ y = 0 \end{cases}$$

VI.34. $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$; generatriz rectilínea

$$x = a\,[\cos\theta \pm (\operatorname{sen}\theta/c)z]$$

$$y = b\,[\operatorname{sen}\theta \mp (\cos\theta/c)z]$$

proyección: $(x/a)\cos\theta + (y/b)\operatorname{sen}\theta = 1$, $z = 0$, que es tangente a la elipse de garganta en

$$x = a\cos\theta, \quad y = b\operatorname{sen}\theta, \quad z = 0$$

VI.35. Generatriz rectilínea

$$x = \sqrt{p}(\mu + z/2\mu) \quad , \quad y = \pm\sqrt{q}(\mu - z/2\mu)$$

proyección $x = \sqrt{p}(\mu + z/2\mu)$, $y = 0$, que es tangente a la parábola $2z = x^2/p$, $y = 0$ en el punto $x_0 = 2\mu\sqrt{p}$, $y_0 = 0$, $z_0 = 2\mu^2$.

VI.36. Generatrices

$$x/\sqrt{p} \pm y/\sqrt{q} = 2\lambda \quad , \quad x/\sqrt{p} \mp y/\sqrt{q} = z/\lambda$$

(λ parámetro);

$$P(0, \pm 2\lambda\sqrt{q}, -2\lambda^2) \quad , \quad Q(2\lambda\sqrt{p}, 0, 2\lambda^2)$$

$$X(\lambda\sqrt{p}, \pm\lambda\sqrt{q}, 0)$$

Lugar: recta $x\sqrt{q} \pm y\sqrt{p} = 0$, $z = 0$.

VI.37. *a*) $\det A > 0$ y $\operatorname{sig} M = \{2, 1\}$, elipse real.
b) $\det A < 0$ y $\det M \neq 0$, hipérbola.
c) $\det A = 0$ y $\det M \neq 0$, parábola.

VI.38. *a*) $\det A > 0$ y $\det M = 0$, dos rectas imaginarias concurrentes: *b*) $\det A = 0$ y $\det M \neq 0$, parábola; *c*) $\det A < 0$ y $\det M = 0$, dos rectas reales concurrentes; *d*) $\det A > 0$ y $\operatorname{sig} M = \{3, 0\}$, elipse imaginaria.

VI.39. $\det A = a$, $\det M = a(a - 4b^2)$; diagonalización de M: $(1, a, a - 4b^2)$. Si $a = 0$, dos rectas paralelas; si $a < 0$, es $\det A < 0$ y $\det M \neq 0$, hipérbola; si $4b^2 < a$, $\det A > 0$ y $\operatorname{sig} M = \{3, 0\}$, elipse imaginaria; si $4b^2 = a > 0$, $\det A > 0$ y $\det M = 0$, dos rectas imaginarias conjugadas; si $0 < a < 4b^2$, $\det A > 0$ y $\operatorname{sig} M = \{2, 1\}$, elipse real.

VI.40. $\det A = 150$, $\lambda_1 = 10$, $\lambda_2 = 15$, $\det M = -3.750$; elipse real; $2x^2 + 3y^2 = 5$; centro $(2, 1)$; direcciones de los ejes $\bar{u}_1(1, -2)$ y $\bar{u}_2(2, 1)$.

VI.41. $\det A = 0$, $\lambda_1 = 0$, $\lambda_2 = 5$, $\det M = -100$; parábola; $\sqrt{5}y^2 = 4x$; dirección de la parábola $\bar{\delta}(1, -2)$; eje (polar de $(0.\ 2,\ 1)$) $2x + y = 0$; vértice $(1, -2)$; tangente en el vértice $x - 2y = 5$.

VI.42. $\det A = \cos^2\theta - 1$, $\lambda_1 = \cos\theta + 1$, $\lambda_2 = \cos\theta - 1$, $\det M = -2(1 + \cos\theta)$; hipérbola; $(\cos^2\theta - 1)x^2 + (\cos\theta - 1)^2y^2 = 2$; centro $[1 : (1 - \cos\theta), 1 : (1 - \cos\theta)]$; direcciones de los ejes $(1, 1)$ y $(1, -1)$.

VI.43. Origen en el foco, «eje x» con la dirección dada; directriz $x\cos\theta + y\operatorname{sen}\theta = p$; parábola

$$x^2\operatorname{sen}^2\theta + y^2\cos^2\theta - 2xy\operatorname{sen}\theta\cos\theta +$$

$$+ 2p(x\cos\theta + y\operatorname{sen}\theta) = p^2$$

la polar de $(0, 1, 0)$ es

$$x\operatorname{sen}^2\theta - y\operatorname{sen}\theta\cos\theta + p\cos\theta = 0$$

eliminando θ, queda $p^2(x^2 + y^2) = 4y^4$.

VI.44. Invariantes métricos:

$$\alpha + \beta = -\det A(a_{11} + a_{12}) : (\det M)$$

y

$$\alpha\beta = (\det A)^3 : (\det M)^2,$$

que son la suma y el producto de las raíces de la ecuación del enunciado.

VI.45. Tomando origen la F y con «eje x» perpendicular a la directriz ($x = d$), la cónica es

$$x^2 + y^2 = e^2(x - d)^2$$

Ecuación de las tangentes desde el origen $x^2 + y^2 = 0$, o sea $y = \pm ix$.

VI.46. La ecuación sólo difiere de la dada en el término independiente, que llamaremos a. La cónica $C \cap \pi$ debe ser degenerada, lo que equivale a que lo sea su proyección sobre $z = 0$, que es

$$-7x^2 + 8xy - 3y^2 + 10y + a = 0$$

luego $a = -35$.

VI.47. Diagonalización de M: $(1, p, -p, a - 1)$.

1. rang $M = 2$, $p = 0$, $x = y$ y $x + y + 2 = 0$.
2. sig $M = (2, 2)$, $a < 1$.
3. Están en el plano tangente $y + z = 0$ y son

$$x = -1 + \lambda(p - 1)$$
$$y = -1 \pm \lambda\sqrt{1 - p}$$
$$z = 1 \mp \lambda\sqrt{1 - p} \quad (\lambda \in \mathbb{R})$$

VI.48. 1. El plano tangente en P_1: $2x - 4y - z + 6 = 0$.
2. El cono circunscrito desde

$$P_2 : 2x^2 + 7y^2 - 3z^2 - 2xz - 4y - 24y + 28 = 0$$

el plano es el polar de $P_2 : 2y = 3$.
3. El punto que tiene a $z = 2$ como plano polar es el $Q(1, -2, 2)$; como Q es de $z = 2$, éste es tangente a C en Q.
4. La intersección de C con $x = h$ debe su degenerada, lo que equivale a $7h^2 + 42 = 0$, que no es posible.

VI.49.

$$U = M\begin{bmatrix}1\\X\end{bmatrix} \quad \text{luego} \quad \begin{bmatrix}1\\X\end{bmatrix} = M^{-1}U$$

El plano U es tangente sii su polo le pertenece, esto es, sii

$$U^t\begin{bmatrix}1\\X\end{bmatrix} = O$$

o sea

$$U^tM^{-1}U = O \quad \text{(ecuación tangencial)}$$

VI.50. Eliminando a y b de $x + (a/2)z = 1$, $(a/2)x + (b/2)y - z + 2 = 0$, $y + (a/2)z = 4$, se tiene

$$x^2 + y^2 + z^2 - x - 4y - 2z = 0$$

VI.51. *a*) sig $M = \{2. 1\}$ y sig $A = \{2, 1\}$; cono real.
b) sig $M = \{3. 1\}$ y sig $A = \{2, 0\}$; paraboloide elíptico.
c) sig M = $\{2. 2\}$ y sig $A = \{2, 1\}$; hiperboloide reglado.

VI.52. *a*) sig $M = \{2. 1\}$ y sig $A = \{1, 0\}$; cilindro parabólico.
b) sig $M = \{3. 1\}$ y sig $A = \{3, 0\}$; elipsoide real.
c) sig $M = \{2. 2\}$ y sig $A = \{1, 1\}$; paraboloide reglado.

VI.53. Diagonalizaciones de M y A: $(1, 1, -1, 3 - \alpha^2)$ y $(1, 1, 1 - \alpha^2)$. $|\alpha| < 1 \rightarrow$ sig $M = \{3, 1\}$ y sig $A = \{3, 0\}$, elipsoide real; $|\alpha| = 1 \rightarrow$ sig $M = \{3, 1\}$ y sig $A = \{2, 0\}$, paraboloide elíptico; $1 < |\alpha| < \sqrt{3} \rightarrow$ sig $M = \{3, 1\}$ y sig $A = \{2, 1\}$, hiperboloide elíptico; $|\alpha| = \sqrt{3} \rightarrow$ sig $M = \{2, 1\}$ y sig $A = \{2, 1\}$, cono real; $|\alpha| > 3 \rightarrow$ sig $M = \{2, 2\}$ y sig $A = \{2, 1\}$, hiperboloide reglado.

VI.54. Diagonalizaciones de M y A: $(1, 1, 3p - 2, p - 2)$ y $(1, 2 - p, p)$. $p < 0$, hiperboloide reglado; $p = 0$, paraboloide reglado; $0 < p < 2/3$, hiperboloide reglado; $p = 2/3$, cono real; $2/3 < p < 2$, hiperboloide elíptico; $p = 2$, cilindro imaginario; $p > 2$, elipsoide imaginario.

VI.55. $a > 0$ y $b \neq 0 \rightarrow$ sig $M = \{3, 1\}$ y sig $A = \{3, 0\}$, elipsoide real; $a < 0$ y $b \neq 0 \rightarrow$ sig $M = \{3, 1\}$ y sig $A = \{2, 1\}$. hiperboloide elíptico; $a > 0$ y $b = 0 \rightarrow$ sig $M = \{3, 0\}$ y sig $A = \{3, 0\}$, cono imaginario; $a < 0$ y $b = 0 \rightarrow$ sig $M = \{2, 1\}$ y sig $A = \{2, 1\}$. cono real; $a = 0$ y $b \neq 0 \rightarrow$ $\rightarrow$ sig $M = \{3, 1\}$ y sig $A = \{2, 0\}$, paraboloide elíptico; $a = 0$ y $b = 0 \rightarrow$ sig $M = \{2, 0\}$ y sig $A = \{2, 0\}$, par de planos imaginarios (con intersección real).

VI.56. $\det M = -32$, $\det A = 32$, $\lambda_1 = \lambda_2 = 2$, $\lambda_3 = 8$; $2x^2 + 2y^2 + 8z^2 = 1$; es de revolución ($\lambda_1 = \lambda_2$). Centro $(-1, 2, 0)$; eje de revolución $\bar{\delta}(2, -1, -1)$; todo vector ortogonal a $\bar{\delta}$ es eje. Elipsoide real.

VI.57. Diagonalizaciones de M y A:

$$(1, 1, 1-\alpha, \alpha - 1/2) \quad \text{y} \quad (1, 1-\alpha, \alpha - 1/2)$$

Para $\alpha < 1/2$ y $\alpha > 1$ hiperboloide elíptico; para $\alpha = 1/2$ y $\alpha = 1$, cilindro imaginario; para $1/2 < \alpha < 1$, elipsoide imaginario. Autovalores de A, $\lambda_1 = 1 - \alpha$, $\lambda_2 = 2\alpha - 1$, $\lambda_3 = 1$; es de revolución para $\alpha = 0$ (para $\alpha = 1$ y $\alpha = 2/3$ imaginaria). Para $\alpha = 3$, $\det M = -10$, $\det A = -10$, $2x^2 - y^2 - z^2 = 1$; centro $(1, 0, 1)$; direcciones de los ejes $(0, 1, 0)$, $(1, 0, 1)$ y $(1, 0, -1)$.

VI.58. *a*) Diagonalizaciones de M y A: $(1, 1, p, p-2)$ y $(1, 1, p-1)$; si $p < 0$, hiperboloide reglado; si $p = 0$, cono real; si $0 < p < 1$, hiperboloide elíptico; si $p = 1$, paraboloide elíptico; si $1 < p < 2$, elipsoide real; si $p = 2$, cono imaginario; si $p > 2$, elipsoide imaginario.

b) $\det M = 12p(p-2)$, $\det A = 24(p-1)$, $\lambda_1 = 2$, $\lambda_2 = 6$, $\lambda_3 = 2(p-1)$; para $p = -1$, $3x^2 + y^2 - 2z^2 = 3/8$.

c) Paraboloide si $p = 1$: $z = x^2 + 3y^2$.

d) Dirección del paraboloide $\bar{\delta}(0, 1, 0)$, direcciones seudoejes $\bar{u}_1(1, 0, -1)$ y $\bar{u}_2(1, 0, 1)$; eje: $x = z = 0$ y vértice $(0, -1/2, 0)$.

VI.59. *a*) $y^2 - 2x + z(ax + by + cz + d) = 0$ con $a = 0$, ya que $\det A = 0$. *b*) Los autovalores no nulos de A iguales, luego $(c-1)^2 = b^2$, o sea $c = 1$ y $b = 0$, con lo que: $y^2 - 2x + z^2 + dz = 0$. *c*) Eje $y = 0$, $z = -d/2$; lugar: $z^2 + 2x = 0$, $y = 0$.

APÉNDICES

CUANTO aquí se va a decir, que quizá debiera haberse dicho al comienzo de este libro, puede considerarse, por extraño que pueda parecer, o como imprescindible o como superfluo. Todo depende de lo que al lector se le suponga conocido y de cómo éste conozca, realmente, lo que se supone que conoce. Con cuantas excepciones fueran precisas, convendría usar estos apéndices para recordar y afianzar aquellos conceptos o resultados que estén olvidados, que estén mal aprendidos o que, incluso, poco menos que se desconozcan, pues de todo ha de haber.

0. Introducción.
1. Nociones sobre lógica.
2. Algo acerca de la teoría intuitiva de conjuntos.
3. Aplicaciones o funciones.
4. Relaciones de equivalencia.
5. Relaciones de orden.
6. Cardinal de un conjunto.
7. Estructuras algebraicas.
8. Grupos.
9. Anillos.
10. Cuerpos.
11. Los ángulos y su medida.
12. Notas, observaciones y complementos.

APENDICE

0

Introducción

Es obligado que, al menos aquí, al comenzar estos apéndices, digamos algo, aunque sea poco, acerca de lo que es la Matemática, de los entes u objetos que en ella se manejan, de los axiomas, de los teoremas y de otras cuestiones relacionadas con todo ello. Vamos a intentarlo, procurando ser concisos y, al tiempo, precisos.

OBJETOS MATEMATICOS Y AXIOMAS. TEOREMAS

[344]

La Matemática se nos presenta hoy como la ciencia que estudia las relaciones entre ciertos entes abstractos, que llamaremos *objetos matemáticos*, caracterizados por satisfacer a unas determinadas propiedades, llamadas *axiomas*, que se aceptan sin demostrar. Las teorías matemáticas están constituidas por *teoremas*, que son proposiciones relativas a los objetos matemáticos obtenidas por *deducción lógica* a partir de los axiomas.

Una conjetura que sea fruto de la intuición sólo será admitida como verdadera cuando llegue a comprobarse recurriendo al razonamiento lógico. Hemos de señalar que la lógica, por si sola, difícilmente puede crear algo nuevo de interés; para edificar una ciencia, se precisa de algo más que de la lógica, se precisa de algún tipo de intuición, que será la guía del proceso creativo(*).

(*) N. Bourbaki: «menos que nunca la matemática se reduce actualmente a un juego puramente mecánico de fórmulas aisladas; más que nunca la intuición reina soberanamente en la génesis de los descubrimientos» (*Las grandes corrientes del pensamiento matemático*).

Lo pertinente es, ahora, dar algunas explicaciones, comentarios y aclaraciones sobre lo que se acaba de decir:

- No se puede introducir el rigor en los razonamientos si se razona sobre objetos definidos sin rigor. De ahí que haya que empezar estableciendo estos objetos matemáticos de manera satisfactoria. Para ello, no se dispone de una fuente de conocimiento distinta de la lógica, no se puede recurrir aquí a una «intuición matemática»; tampoco es dado acudir al mundo real, al mundo físico, aun cuando dichos objetos pueden proceder de él mediante procesos

de abstracción. Así pues, se renuncia a la fundamentación metafísica de la matemática; se abandona el empeño, de tiempos pasados, de fijar qué cosas son los objetos matemáticos. En lugar de ello, se parte de axiomas que establecen las propiedades de dichos objetos: desde el punto de vista matemático, los axiomas definen implícitamente los objetos a los que se refieren (véase [420], NOTA 4, al final de los Apéndices).

- En toda teoría matemática, se parte de ciertos elementos u objetos cuya existencia y propiedades se admiten y se fijan mediante un sistema de axiomas. Un axioma es una proposición que se acepta y, en consecuencia, no precisa demostración. Nótese que, en matemáticas, los axiomas no son «propiedades tan claras y evidentes que no necesitan demostración»; aquí, los axiomas son ciertos por definición, esto es, se admite su veracidad, se suponen verdaderos.
- A cada proposición, o aseveración relativa a los objetos matemáticos, se la tratará de asignar uno de los términos verdadero (v) o falso (f), que se llamarán valores de verdad, de acuerdo con las reglas de la lógica, que posteriormente se esbozarán. Los axiomas de una teoría son, por definición, proposiciones verdaderas de la misma; las demás proposiciones verdaderas son los teoremas.
- A los axiomas de una teoría matemática se les exige que, en ella, no se encuentre ninguna contradicción, esto es, ninguna proposición que sea verdadera y falsa a la vez, lo que se expresará diciendo que los axiomas han de ser compatibles, coherentes o no contradictorios. Una teoría se aceptará, pues, en tanto que sus axiomas resulten ser compatibles; es más, la ausencia de contradicciones, y no otra cosa, determinará la validez de la teoría (véase [418], NOTA 2). La compatibilidad de un sistema de axiomas es lo que garantiza, matemáticamente hablando, la existencia de los objetos matemáticos que los axiomas definen; no es aceptable recurrir a ningún otro criterio para establecer la existencia de tales objetos.
- Si, en una teoría matemática, una proposición no es ni verdadera ni falsa, se dirá que es una proposición indecidible; esto es, las proposiciones indecidibles son aquellas que no pueden ser demostradas ni refutadas con los axiomas de la teoría que se considera. Una teoría se dice que es completa si en ella no hay proposiciones indecidibles (véase [419], NOTA 3). Aunque en las teorías matemáticas son escasas las proposiciones indecidibles, los axiomas de aquéllas no son tan restrictivos como para evitar que éstas existan. Salvo las posibles proposiciones indecidibles y bajo axiomas no contradictorios, una proposición será, pues, verdadera si no es falsa y será falsa si no es verdadera.

MODELOS MATEMATICOS

Recordando algo que ya venimos diciendo, es de señalar que, en las teorías matemáticas, los objetos y los axiomas que las gobiernan no han de ser, necesariamente, idealizaciones de objetos sensibles y de las propiedades que éstos pudieran tener. No obstante, sí pudieran serlo y, de hecho, en muchas ocasiones lo son; cuando así ocurra, lo será del siguiente modo:

[345]

Tras analizar un problema del mundo real y percibir las relaciones que existen entre sus distintos elementos, un proceso de abstracción puede conducir a la creación de un *modelo matemático* del problema. Como resultado de este proceso, se definen unos objetos matemáticos (que no existen en la naturaleza; son idealizaciones de objetos naturales) que han de satisfacer a ciertas relaciones (que son reflejo de las propiedades que supuestamente verifican los objetos naturales); de este modo se establecen unos axiomas, que serán el punto de partida de la correspondiente teoría matemática. Basándose en los axiomas y por deducción lógica se obtienen unas consecuencias que son los teoremas de la teoría. Finalmente, las consecuencias se aplican a la resolución del problema de partida.

A este respecto, hay que observar que, en un modelo matemático, los teoremas son consecuencias lógicas de los axiomas y no deben interpretarse como relaciones reales concernientes al problema que dio origen al modelo. La validez de una teoría no depende de su aplicabilidad al problema. La mayor o menor adecuación de los resultados a la situación real estará relacionada, en todo caso, con lo atinada que haya sido la elección de los axiomas, esto es, con el mayor o menor acomodo de éstos a la situación concreta de la que proceden.

APENDICE

1

Nociones sobre lógica

En matemáticas, la lógica es la ciencia de las deducciones, de la demostración; en ella se estudian las relaciones que existen entre las proposiciones, con independencia del significado concreto que éstas puedan tener y de su «valor de verdad» (esto es, de si son verdaderas o falsas). En los tratados de lógica, en los que estos temas se estudian con mucho mayor detalle y precisión de lo que aquí vamos a necesitar, dicen «sea p una "variable de enunciado"» o «sea p una "variable proposicional"» como equivalente de «sea p una proposición o un enunciado arbitrario, no especificado, que puede ser verdadero o puede ser falso». Nosotros nos contentaremos con decir «sea p una proposición cualquiera» o solamente «sea p una proposición».

PROPOSICIONES O ENUNCIADOS

[346]

Una proposición (o enunciado) es cualquier frase, o expresión escrita, de la que puede decirse que es verdadera o que es falsa, pero no ambas cosas a la vez, es decir, a la que se puede asignar uno y sólo uno de los «valores de verdad» v (de lo verdadero) o f (de lo falso).
Una proposición o enunciado simple consta de un verbo, un sujeto y un predicado; en ella se dice algo del sujeto; lo que se dice del sujeto es el predicado.

EJEMPLOS

Para ilustrar lo anterior, vamos a considerar unas frases que son proposiciones y otras que no lo son:

- «Séneca fue un filósofo estoico» es una proposición verdadera.
- «El acido sulfúrico es una golosina» es una proposición falsa.
- «Colón no nació en Génova» es una proposición; seguro que o es verdadera o es falsa, aunque no sepamos, hoy, cual de ellas es.
- «El trabajo ennoblece» no es una proposición; unas personas lo consideran verdadero y otras estiman que es falso.
- «¡Viva tu madre, morena!» no es una proposición; no tiene sentido asignar aquí valor verdadero ni valor falso.
- Esta proposición no es cierta» no es una proposición, pues ni puede ser verdadera ni puede ser falsa(*).

(*) Esta es la «paradoja del mentiroso» o de Epiménides. Si fuese verdadera, entonces sería cierto que no es cierta, luego sería falsa. Si fuese falsa, entonces sería falso que es falsa, luego sería verdadera.

OBSERVACION

En matemáticas, la veracidad o falsedad de una proposición depende de las hipótesis de partida, o sea, de la teoría matemática que se esté considerando. Así, la proposición «el producto es una operación conmutativa» es cierta si se refiere al producto de número reales, pero es falsa si lo que se considera es el producto de matrices.

La proposición «la suma de las medidas de los tres ángulos de un triángulo es igual a 180°» es cierta en la Geometría Euclídea, esto es, en aquella en la que se admite el postulado de las paralelas (véase [420], NOTA 4b); esta proposición es falsa en las geometrías en las que se sustituye dicho postulado por otro que le contradiga, como las de Lobachewski y de Riemann.

OPERACIONES LOGICAS

Al combinar proposiciones conforme a las reglas de la lógica, esto es, al reunir proposiciones y ciertos signos lógicos, se obtienen nuevas proposiciones. Estos resultados, de operar lógicamente con proposiciones, quedan determinados por los valores de verdad que se les asigna, los cuales dependerán (serán función) de los que tengan las proposiciones de partida.

[347]

Al preceder a una proposición p del signo «no», se obtiene la *negación* de p, que es verdadera si p es falsa y recíprocamente. En los tratados de lógica, a la negación «no p» se la representa poniendo «$\neg p$», «$\sim p$» o «$\bar{p}$».

A partir de dos proposiciones p y q y recurriendo a un signo $*$, se forma la nueva proposición $p * q$. A los signos $*$ se les llama *conectivas* (u operadores lógicos); cada una de ellas queda caracterizada por el valor de verdad que se asigna a $p * q$ en función de los de p y de q. Aquí estamos interesados en las cuatro conectivas «o», «y», «$\rightarrow$» y «$\leftrightarrow$»:

		Valores de verdad			
p y q son dos proposiciones dadas; en función de ellas, se definen ⌣	p	v	v	f	f
	q	v	f	v	f
disyunción: «p o q» (también se pone $p \vee q$)		v	v	v	f
conjunción: «p y q» (también se pone $p \wedge q$)		v	f	f	f
condicional: «$p \rightarrow q$» (se lee: si p entonces q)		v	f	v	v
bicondicional: «$p \leftrightarrow q$» (se lee: p si y sólo q)		v	f	f	v

EJERCICIO

Sean p y q dos proposiciones cualesquiera y considérense las proposiciones:

a) $(no\,q) \rightarrow (no\,p)$.
b) $(p\,y\,q)$ o $[(no\,p)\,y\,(no\,q)]$.

Hallar los valores de verdad de *a*) y de *b*). en función de los de p y de q.

RESOLUCION

Recurriendo a las definiciones de «*no*», «*o*», «*y*» y «$\rightarrow$», se obtiene fácilmente el siguiente cuadro:

p	q	$no\,p$	$no\,q$	*a*)	$(no\,p)\,y\,(no\,q)$	$p\,y\,q$	*b*)
v	v	f	f	v	f	v	v
v	f	f	v	f	f	f	f
f	v	v	f	v	f	f	f
f	f	v	v	v	v	f	v

OBSERVACIONES

1. Hay dos usos diferentes de la conjunción disyuntiva «*o*». Hay una «*o*» exclusiva, como en «te portas bien o te castigo» (si se porta bien no hay castigo); en este caso, decir «p *o* q» significaría «p o q pero no ambos». Hay también una «*o*» inclusiva, como en «para viajar en tren o en autobús hay que pagar billete»; en este caso, decir «p *o* q» significa «p *o* q o ambos». En la anterior definición de disyunción, a la partícula «*o*» se le da el significado inclusivo; en ella, «*o*» tiene el significado de «una al menos de las dos».
2. Los valores de verdad que se han atribuido a la condicional «$p \rightarrow q$» no responden a lo que en el lenguaje usual significa «si p, entonces q», a pesar de que de este modo se suele leer la condicional. A este respecto, nótese que «$p \rightarrow q$» es falsa sólo si p es verdadera y q es falsa; así, por ejemplo, las tres condicionales siguientes son verdaderas: «si $1 + 1 = 2$, entonces por dos puntos distintos pasa una sola recta» (a pesar de que $1 + 1 = 2$ no es causa de que por dos puntos distintos pase una sola recta), «si $1 + 1 = 3$, entonces $1 + 1 = 2$» (a pesar de que $1 + 1 = 3$ sea falsa) y «si $1 + 1 = 3$, entonces $1 + 1 = 4$» (donde $1 + 1 = 3$ y $1 + 1 = 4$ son falsas). La condicional «$p \rightarrow q$» no debe leerse «p implica q»: a esta expresión se la atribuirá más adelante un significado distinto.

[348]

Dadas varias proposiciones p, q, r, ... y recurriendo a las conectivas (*no*, *o*, *y*, $\rightarrow$, $\leftrightarrow$) se pueden construir proposiciones cada vez más complejas. Estas reuniones de proposiciones, agrupadas mediante conectivas, se llaman *expresiones lógicas*. Acudiendo a la anterior tabla de [347], que da los valores de verdad de las distintas conectivas, se pueden determinar los valores de verdad de una expresión lógica, sin más que hallar, en función de los de p, q, r, ..., los de las sucesivas proposiciones parciales de que conste la expresión lógica dada.

EJEMPLO

Los valores de verdad de la expresión lógica

$$(p \text{ o } r) \rightarrow [(no\ q) \text{ y } r]$$

(donde p, q y r son tres proposiciones) se obtienen como se señala en la siguiente tabla:

p	v	v	v	v	f	f	f	f
q	v	v	f	f	v	v	f	f
r	v	f	v	f	v	f	v	f
$p \text{ o } r$	v	v	v	v	v	f	v	f
$no\ q$	f	f	v	v	f	f	v	v
$(no\ q) \text{ y } r$	f	f	v	f	f	f	v	f
$(p \text{ o } r) \rightarrow [(no\ q) \text{ y } r]$	f	f	v	f	f	v	v	v

EJERCICIO

Se ha producido un robo y se sospecha que un tal Dimas fue el ladrón. Se sabe que:

a) Si no llovía (cuando se produjo el robo), entonces Dimas fue el ladrón.
b) O bien, llovía y Dimas fue el ladrón, o bien, el ladrón usó guantes.
c) Si Dimas no fue el ladrón, entonces no llovía y el ladrón usó guantes.

¿Puede asegurarse que Dimas fue el ladrón?

RESOLUCION

Llamemos p, q y r a las proposiciones:

$p \equiv$ «durante el robo llovía»
$q \equiv$ «Dimas fue el ladrón»
$r \equiv$ «el ladrón usó guantes»

Según el enunciado, se sabe que las tres proposiciones a, b y c, siguientes son verdaderas:

$$a \equiv (no\, p) \rightarrow q \quad ; \quad b \equiv (p\, y\, q)\, o\, r \quad ; \quad c \equiv (no\, q) \rightarrow [(no\, q)\, y\, r]$$

Construyendo la «tabla de verdad» de las proposiciones a, b y c, se obtiene:

	1.ª	2.ª	3.ª	4.ª	5.ª	6.ª	7.ª	8.ª
p	v	v	v	v	f	f	f	f
q	v	v	f	f	v	v	f	f
r	v	f	v	f	v	f	v	f
a	v	v	v	v	v	v	f	f
b	v	v	v	f	v	f	v	f
c	v	v	f	f	v	v	v	f

Como a, b y c son verdaderas y esto sólo se produce en las columnas 1.ª, 2.ª y 5.ª, resulta que los valores de verdad de p, q y r han de ser los de una de estas tres columnas y en las tres el valor de verdad de q es v, luego q es verdadera, o sea, Dimas fue el ladrón.

CUANTIFICADORES

[349]

Si $p(x)$ es una expresión que se convierte en proposición al sustituir la letra x por un objeto matemático, se dice que p es una *función proposicional*. También hay funciones proposicionales con más de una letra.

Si $p(x)$ es siempre una proposición verdadera, para cualquiera que sea el objeto matemático que sustituya a x, entonces se pondrá

$$\forall x,\ p(x)$$

Esta proposición se lee «todo x satisface a $p(x)$» o «para todo objeto matemático se verifica p». El signo lógico $\forall$ se llama *cuantificador universal.*

Si $p(x)$ es alguna vez una proposición verdadera, al sustituir x por al menos un cierto objeto matemático, entonces se pondrá

$$\exists x,\ p(x)$$

Esta proposición se lee: «existe algún x que satisface a $p(x)$» o «existe al menos un objeto matemático para el que se verifica p». El signo $\exists$ se llama *cuantificador existencial.*

Las negaciones de $\forall x,\ p(x)$ y de $\exists x,\ p(x)$ son las proposiciones:

$$no\,[\forall x,\ p(x)] \equiv \exists x,\ no\, p(x) \quad ; \quad no\,[\exists x,\ p(x)] = \forall x,\ no\, p(x)$$

EJEMPLOS

1. Si x denota a un número real cualquiera, la función proposicional $p(x) \equiv «x^3 > x^2»$ es una proposición verdadera para $x = 2$ y es una proposición falsa para $x = 1/2$.
2. Si x denota a un triángulo cualquiera, consideremos la función proposicional $p(x) \equiv$ «el baricentro de x (punto en el que se cortan sus medianas) coincide con el circuncentro de x (punto en el que se cortan sus mediatrices). La proposición $\exists x,\ p(x)$ es cierta ya que $p(x)$ es verdadera cuando x es un triángulo equilátero, aunque no lo sea para los demás triángulos.
3. Si x denota a un ángulo cualquiera, considérese la función proposicional $p(x) \equiv «\operatorname{sen} 2x = 2 \operatorname{sen} x \cos x»$. La proposición $\forall x,\ p(x)$ es verdadera, ya que para cualquiera que sea el ángulo x, los números $\operatorname{sen} 2x$ y $2 \operatorname{sen} x \cos x$ son iguales.
4. La negación de «todo número impar es primo» es «hay algún número impar que no es primo». La negación de «existe algún número real que es menor que su parte entera» es «todo número real es mayor o igual que su parte entera».

EJERCICIO

Epiménides de Cnosos (siglo VI a. de C.) decía «Todos los cretenses son mentirosos y yo soy cretense, luego miento». Alguien a la vista de ello, razona como sigue:

Si Epiménides mintió en lo que dijo, entonces los cretenses no eran mentirosos, luego Epiménides, por ser cretense, no era mentiroso y, consecuentemente, no mintió en lo que dijo. Se llega así, pues, a una contradicción. ¿Este razonamiento es correcto?

RESOLUCION

En el razonamiento se ha cometido el error de decir que la negación de «los cretenses (todos) son mentirosos» es «los cretenses (todos) no son mentirosos», cuando tal negación es «algún cretense (uno al menos) no es mentiroso». Dicho de otra forma,

llamando: $p(x) \equiv$ «el cretense x es mentiroso»

se tiene que: $\forall x,\ p(x) \equiv$ todos los cretenses son mentirosos

y también: $no\,[\forall x,\ p(x)] \equiv \exists x,\ no\,p(x) \equiv$ algún cretense no es mentiroso

Con independencia de lo anterior, conviene notar que:

$\forall x,\ no\,p(x) \equiv$ los cretenses (todos) no son mentirosos

$no\,[\forall x,\ no\,p(x)] \equiv \exists x,\ p(x) \equiv$ algún cretense es mentiroso.

IMPLICACIONES Y TEOREMAS. DEMOSTRACIONES

[350]

Se llamará *implicación* a toda condicional $p \to q$ verdadera en la que p se supone verdadera; también se suele pedir que existe una relación de causa-efecto entre p y q. Si $p \to q$ es una implicación, se la expresa poniendo $p \Rightarrow q$; a la proposición p se la llama premisa o *hipótesis* y la q recibe el nombre de conclusión o *tesis*. La implicación $p \Rightarrow q$ se lee de las siguientes formas «p *implica* q», «p es *condición suficiente para* q» y «q es *condición necesaria para* p».

En una teoría matemática, se llaman *teoremas* a las proposiciones verdaderas relativas a los objetos matemáticos de la misma. Los términos teorema e implicación se suelen considerar equivalentes(*).

Si $p \to q$ y $q \to p$ son implicaciones, se pondrá $p \Leftrightarrow q$ y se dirá que «p es *equivalente* a q», que «p es *condición necesaria y suficiente para* q» o que «p *si y sólo si* q» (esto último se suele abreviar poniendo «p sii q»).

OBSERVACIONES

1. Se entiende que (en una implicación $p \Rightarrow q$) hay una relación de causa-efecto entre p y q si en alguno de los casos en los que p no se verifica tampoco se verifica q. Así, en la implicación, «si n es un número par, entonces n no es primo», hay una relación de causa-efecto entre $p \equiv$ «n es número par» y $q \equiv$ «n no es primo», ya que $n = 5$ (que no es par) es primo. En la proposición $p \equiv$ «en la geometría euclídea, si un triángulo es equilátero, entonces sus ángulos suman 180°» (que es una implicación verdadera), no hay relación de causa-efecto entre $p \equiv$ «el triángulo es equilátero» y $q \equiv$ «los ángulos del triángulo suman 180°», ya que q es cierta para cualquier triángulo, sea o no equilátero.
2. Nótese que al hablar de lo que es una implicación $p \Rightarrow q$, se ha dicho que «p se supone verdadera» y no se dijo «p es verdadera». Aceptando, pues, que la premisa p es cierta, la implicación se prueba al comprobar la relación de causa-efecto entre p y q. Una teoría es válida a causa de que sus teorema o implicaciones son proposiciones verdaderas, con independencia de la veracidad de sus axiomas. A este respecto, considérense las proposiciones $p \equiv$ «el número -1 admite una raíz cuadrada que es real y positiva» y $q \equiv$ «el número -4 admite una raíz cuadrada que es real y positiva»; en este supuesto, $p \to q$ es una implicación (proposición verdadera) ya que, suponiendo que p es verdadera (sea x la raíz cuadrada que se cita en p), entonces también lo sería q (la raíz que se cita en q sería $2x$) y ello como consecuencia de p.

(*) Véase [423], NOTA 7.

[351]

El proceso lógico que, a partir de una proposición ya probada p, conduce a la veracidad de otra proposición q, es la *demostración* del teorema $p \Rightarrow q$. La demostración directa se apoya en que, de ser ciertas p y $p \Rightarrow q$, entonces también lo es q. La demostración por el *contrarrecíproco* consiste en probar que $no\,q \Rightarrow no\,p$, en lugar de probar $p \Rightarrow q$. La demostración por *reducción al absurdo* consiste en suponer que q es falsa (con lo que la hipótesis pasa a ser $p\,y\,(no\,q)$ y probar que de ello se deduce alguna contradicción.

OBSERVACIONES

- Generalmente las demostraciones directas son las más cortas y elegantes; ellas conducen sin rodeos a la tesis del teorema.
- Un teorema $p \Rightarrow q$, que llamaremos *teorema directo*, es equivalente al teorema $no\,q \Rightarrow no\,p$, que se llama *teorema contrarrecíproco* del dado; demostrando uno de ellos, queda entonces demostrado el otro. Los teoremas $q \Rightarrow p$ (*teorema recíproco*) y $no\,p \Rightarrow no\,q$ (*teorema contrario*) son equivalentes entre sí, pero no son equivalentes al $p \Rightarrow q$ (directo).
- Un teorema $p \Rightarrow q$ es equivalente a cualquiera de los teoremas:

$$\begin{array}{lcl} [p\,y\,(no\,q)] & \Rightarrow & (no\,p) \\ [p\,y\,(no\,q)] & \Rightarrow & q \\ [p\,y\,(no\,q)] & \Rightarrow & r \quad (r = \text{proposición falsa}) \end{array}$$

Este hecho es el que da lugar al método de demostración por reducción al absurdo.

[352]

Refutar un supuesto teorema $p \Rightarrow q$ es comprobar que es falso, lo que se indica poniendo $p \not\Rightarrow q$. Para ello, se puede suponer que p y q son ciertas y buscar, a partir de ello, un resultado contradictorio.

También se puede mostrar que $p \not\Rightarrow q$ encontrando un *contraejemplo*, esto es, algún caso concreto en el que p sea verdadera y q sea falsa. Así, para probar que la proposición «$\forall x,\ p(x) \rightarrow q(x)$» es falsa, bastará con localizar un objeto matemático a tal que $p(a)$ sea verdadera y $q(a)$ sea falsa.

APENDICE

2

Algo acerca de la teoría intuitiva de conjuntos

No se nos oculta que en la teoría intuitiva de conjunto se presentan paradojas y contradicciones y que ello ha conducido a la axiomatización de la teoría de conjuntos (véanse [420], NOTA 4 *c* y *d*, y [421], NOTA 5). No obstante, estimamos que éste no es sitio adecuado para abordar el estudio axiomático de los conjuntos; opinamos así por varios motivos. El primero de ellos es que, en nuestro estudio del Algebra Lineal, no nos van a salir al paso las referidas contradicciones de la teoría intuitiva de conjuntos, salvo que nos desviemos de nuestro camino para ir a su encuentro. Pero, además, la extensión y la dificultad de la teoría axiomática de conjuntos hacen poco menos que imposible tratar de incluirla aquí, en un apéndice.

A este respecto, conviene tener en cuenta que, en los asuntos del enseñar y del aprender, hay siempre que decir la verdad. pero, para que sea conveniente decir toda la verdad, es necesario que el que la recibe esté en condiciones de poderla entender.

SOBRE DE LA NOCION INTUITIVA DE CONJUNTO

Según ya se ha indicado, la noción de conjunto se conceptúa como algo primitivo, no definible. Ello no es obstáculo para que incluyamos aquí algunos comentarios y puntualizaciones, al objeto de afianzar el concepto primario de conjunto que al lector se le supone(*).

[353]

«Un *conjunto* es la reunión en un todo de determinados objetos bien definidos y diferenciables entre sí, que se llaman elementos del conjunto». Un conjunto está generado por sus elementos y es un objeto nuevo, distinto de cada uno de los elementos que lo integran; los elementos de un conjunto, distintos entre sí, han de estar determinados antes de que se reúnan para formar el conjunto. Si *a* es un elemento del conjunto *A*, se pondrá $a \in A$; el símbolo $\in$ se llama *signo de pertenencia*.

Dos *conjuntos iguales* son aquellos que tienen los mismos elementos.

(*) En el año 1902, Bertrand Russell encontró el conjunto paradójico que hoy lleva su nombre (véase [421]), que puso en tela de juicio a la teoría intuitiva de conjuntos. En el año 1903 se lo comunicó al también matemático Grottob Frege, cuando éste estaba a punto de publicar el segundo volumen de su gran obra *Leyes básicas de la Aritmética*, que se apoyaba en la teoría intuitiva de conjuntos. En un apéndice de este segundo volumen, Frege insertó la siguiente apostilla: «Difícilmente puede ocurrirle a un escritor científico algo más desgraciado que ver ceder los cimientos de su obra cuando se está terminando. Tal es la situación en que me ha puesto una carta de Mr. Bertrand Russell».

OBSERVACIONES

1. Conviene notar que, en el supuesto de disponer de ciertos elementos y por el hecho de considerarlos agrupados, se conviene en que existe un nuevo objeto, que se llama conjunto formado por los elementos dados.
2. En ocasiones, en lugar de conjunto y con su mismo significado, se utilizan también los términos *familia* o *colección*.
3. Cuando se diga que «a es un elemento genérico del conjunto A» o que «a recorre A», se deberá entender que por a se denota a uno cualquiera de los elementos de A.
4. Si b no es un elemento del conjunto A, se pondrá $b \notin A$.
5. El significado intuitivo de la palabra conjunto es del dominio general; para designar a ciertos conjuntos particulares se dispone de palabras concretas como fauna o cabellera. En geometría se utiliza el término «lugar geométrico» como sinónimo de conjunto.
6. Para la igualdad de conjuntos se verifican las siguientes propiedades: para cualesquiera conjuntos A, B y C, es: a) la proposición $A = A$ es verdadera; b) las proposiciones $A = B$ y $B = A$ son equivalentes; c) si es $A = B$ y $B = C$, entonces es $A = C$. Estas tres propiedades (reflexiva, simétrica y transitiva) serán las que, como se verá más adelante, caractericen a las relaciones de equivalencia.

EJEMPLO

Los conjuntos de los números naturales, de los números enteros, de los números racionales, de los números reales y de los números complejos se denotan, respectivamente por $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$. Si X es un conjunto al que pertenece el número 0 (cero), entonces se pone X^* para designar al conjunto formado por todos los elementos de X menos el cero. Si X es un conjunto formado por números positivos y negativos (por ejemplo, para X igual a $\mathbb{Z}$, $\mathbb{Q}$ o $\mathbb{R}$), X_+ y X_- representarán a los conjuntos que forman los elementos de X que son positivos y negativos, respectivamente.

[354]

Para complementar las anteriores ideas sobre los conjuntos, digamos que:

- Se admite que existe un conjunto $\emptyset$ que carece de elementos, al que se llama *conjunto vacío*.
- Cuando se consideren varios conjuntos y siempre que sea de utilidad, se recurrirá al conjunto U formado por todos aquellos objetos que pertenecen a alguno (a uno, al menos) de los conjuntos dados; a dicho conjunto U se le llama *universal* o *referencial*.
- Para señalar cuáles son los elementos que constituyen un conjunto, se puede proceder: 1) enumerando todos y cada uno de dichos elementos (por *extensión*); o 2) definiendo los elementos mediante propiedades que los caractericen (por *compresión*).
- Un conjunto se denotará encerrando entre llaves a sus elementos, si viene dado por extensión, o a la caracterización de los mismos, si viene dado por compresión.

EJEMPLOS

- Al conjunto que forman las letras a, b, c y d se le denota poniendo $\{a, b, c, d\}$. Al conjunto de los cuadrados de los números naturales se le denota poniendo $\{n^2/n \in \mathbb{N}\}$.
- Los conjuntos $A = \{0, 1, 2, 3, 4\}$ y $B = \{n \in \mathbb{N}/n^2 - 2n - 8 \leq 0\}$ son iguales, ya que la relación $n^2 - 2n - 8 \leq 0$ equivale a la $(n + 2)(n - 4) \leq 0$, esto es, a la $-2 \leq n \leq 4$ y ésta es satisfecha por los números naturales 0, 1, 2, 3 y 4 y sólo por ellos.
- No cuesta nada comprobar que

$$\left\{n \in \mathbb{Z} \Big/ \frac{n^2+3}{n^2+2} \in \mathbb{N}\right\} = \emptyset$$

- Si A y B son dos puntos del espacio ordinario E (espacio euclídeo ordinario, tridimensional), situados, uno del otro, a distancia unidad, el conjunto

$$\{X \in E/\text{el triángulo } ABX \text{ tiene área unidad}\}$$

es la superficie cilíndrica que tiene por eje a la recta AB y cuyo radio es 2.
- Dada una función proposicional p, se llama *conjunto de verdad* de p al conjunto P formado por los objetos matemáticos a tales que $p(a)$ es una proposición verdadera. Así, por ejemplo, si $p(x)$ es «para el número real x se verifica que sen $x = 0$», el conjunto de verdad es $P = \{n\pi/n \in \mathbb{Z}\}$

SUBCONJUNTOS; COMPLEMENTARIO

Vamos a considerar aquí a los conjuntos que se pueden formar tomando elementos de un cierto conjunto dado U, que se toma como referencial fijo; tales conjuntos se llaman subconjuntos del referencial.

[355]

Se dice que un conjunto A es un *subconjunto* de otro conjunto U si todos los elementos de A son elementos de U; en tal caso, también se dice que A es una *parte* de U y que A está *incluido* en U, lo que se expresa poniendo $A \subset U$, donde $\subset$ se llama *signo de inclusión*. Entre los subconjuntos de U, se encuentran el vacío $\emptyset$ y el propio U; si $A \subset U$ y es $A \neq \emptyset$ y $A \neq U$, se dice que A es una parte propia de U.
Al conjunto que tiene por elementos a los subconjuntos de U se le llama *conjunto de las partes* de U, el cual se denota por $\mathcal{P}(U)$. La relación $A \subset U$ equivale a la $A \in \mathcal{P}(U)$.

OBSERVACIONES

- Si $A \subset U$ y $A \neq U$, se dice que A está estrictamente incluido en U. Hay quienes el signo de inclusión lo escriben poniendo $\subseteq$ y reservan el signo $\subset$ para la inclusión estricta.

- La relación $a \in A$, que significa que a es un elemento del conjunto A, equivale a la relación $\{a\} \subset A$, que significa que el conjunto, que tiene a a como único elemento, es un subconjunto de A.
- La inclusión de conjuntos es reflexiva (es decir, para todo conjunto A es $A \subset A$), es antisimétrica (es decir, si es $A \subset B$ y $B \subset A$, entonces es $A = B$) y es transitiva (es decir, si es $A \subset B$ y $B \subset C$, entonces es $C \subset A$).

EJEMPLOS

- Llamando $A = \{x \in \mathbb{R} / x^2 \leq x\}$ y $B = \{x \in \mathbb{R} / x^2 \leq 2x\}$, se verifica que $A \subset B$; nótese que A y B son los intervalos $A = [0, 1]$ y $B = [0, 2]$.
- Para cada uno de los siguientes conjuntos U, el conjunto $\mathcal{P}(U)$ es el que se indica:

Si $U = \{a\}$, entonces $\mathcal{P}(U) = \{\emptyset, \{a\}\}$

Si $U = \{a, b\}$, entonces $\mathcal{P}(U) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$

Si $U = \{a, b, c\}$, entonces $\mathcal{P}(U) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{c, a\}, U\}$

Si $U = \emptyset$, entonces $\mathcal{P}(U) = \{\emptyset\}$ y $\mathcal{P}(\mathcal{P}(U)) = \{\emptyset, \{\emptyset\}\}$

[356]

Si A es un subconjunto del conjunto U, se llama *complementario* de A respecto de U al conjunto formado por los elementos de U que no pertenecen a A y se le denotará por $\complement_U A$ o abreviadamente poniendo $\complement A$ ó $\bar{A}$.

Según esta definición y si A y B son subconjuntos de U, se verifica:

$$\complement_U \emptyset = U; \quad \complement_U U = \emptyset; \quad \complement_U(\complement_U A) = A$$

$$A \subset B \Leftrightarrow \complement_U B \subset \complement_U A$$

Dados dos conjuntos cualesquiera A y B, se llama *diferencia* $A - B$ al conjunto $A - B = \{a \in A / a \notin B\}$.

Las anteriores propiedades se comprueban fácilmente, por lo que se dejan como ejercicios.

EJEMPLOS

- En el conjunto referencial de los números enteros, si A es el conjunto que forman los múltipos de 3, el complementario de A está formado por los números que al dividirlos entre 3 dan de resto 1 ó 2.
- Sea $P = \{x \in U / p(x)$ es una proposición verdadera$\}$, donde p es una cierta función proposicional. El complementario del conjunto P es el conjunto $\bar{P} = \{x \in U / p(x)$ es falsa$\}$.

UNION E INTERSECCION DE CONJUNTOS

[357]

Se llama *unión* de dos conjuntos A y B al conjunto formado por todos los elementos que pertenecen a uno, al menos, de los conjuntos A o B, al cual se le denota por $A \cup B$. Se llama *intersección* de dos conjuntos A y B al conjunto formado por todos los elementos que pertenecen a los dos conjuntos A y B, al cual se denota por $A \cap B$. Se dice que A y B son *disjuntos* si no tienen elementos comunes, o sea, si $A \cap B = \emptyset$.

Dado un conjunto U, se llama *partición* de U a toda familia de subconjuntos de U disjuntos dos a dos y tales que la unión de todos ellos es U.

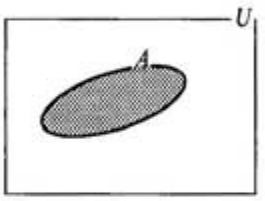


Figura 1.a.

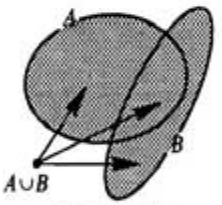


Figura 1.b.

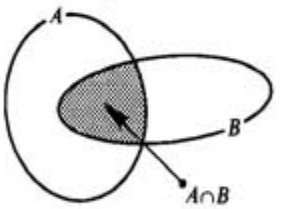


Figura 1.c.

OBSERVACIONES

- Para ayudar a la intuición, suele ser útil representar los conjuntos por regiones del plano; cada conjunto queda simbolizado por los puntos interiores a una curva cerrada simple. Estos esquemas se llaman *diagramas de Venn*. Con ellos suele ser fácil obtener una impresión sobre si son o no ciertas las propiedades (relativas a la unión, la intersección, los complementarios) que se estudien; no obstante, los razonamientos sobre diagramas no tienen valor demostrativo, salvo para refutar, con un contraejemplo, una supuesta propiedad.

  Así, por ejemplo, en la Figura 1.a se representa un conjunto universal U (región interior al rectángulo) y un conjunto A (zona sombreada); la zona no sombreada del rectángulo representa al complementario de A. En la Figura 1.b se representan dos conjuntos A y B y su unión (zona sombreada); la Figura 1.c representa la intersección de dos conjuntos A y B no disjuntos.
- Si A y B son los conjuntos definidos por compresión:

$$A = \{x \in U / p(x)\} \quad \text{y} \quad B = \{x \in U / q(x)\}$$

  donde U es un conjunto universal y p y q son dos funciones proposicionales, entonces:

$$A \cup B = \{x \in U / p(x) \text{ o } q(x)\} \quad ; \quad A \cap B = \{x \in U / p(x) \text{ y } q(x)\}$$

- Se dice que una familia de conjuntos es un *recubrimiento* de un conjunto U si la unión de todos los conjuntos de la familia incluye a U.

EJERCICIOS

Compruébese que, dados dos conjuntos A y B, se verifica que:

1. $A \cup B = A \Leftrightarrow B \subset A$.
2. $A \cap B = A \Leftrightarrow A \subset B$.
3. $A - B = A \cap \bar{B}$.

($\bar{B}$ denota al complementario de B, respecto de cierto universal U).

COMPROBACIONES

1. Esta equivalencia es verdadera ya que $A \cup B = A$ y $B \subset A$ son ambas equivalentes a $(x \in B) \Rightarrow (x \in A)$.
2. Las relaciones $A \cap B = A$ y $A \subset B$ equivalen ambas a $(x \in A) \Rightarrow (x \in B)$, luego son equivalentes entre sí.
3. La igualdad a comprobar se obtiene de:

$$A - B = \{x \in U / x \in A \text{ y } x \notin B\} = \{x \in U / x \in A\} \cap \{x \in U / x \notin B\} = A \cap \bar{B}$$

[358]

Propiedades de las operaciones booleanas(*). Para cualesquiera que sean los conjuntos A, B y C, se verifica que:

1. $A \cup \emptyset = A \qquad A \cap \emptyset = \emptyset$
2. $A \cup A = A \cap A = A$ (idempotencia)
3. $A \cup B = B \cup A \qquad A \cap B = B \cap A$ (conmutatividad)
4. $\left.\begin{array}{l} A \cup (B \cup C) = (A \cup B) \cup C \\ A \cap (B \cap C) = (A \cap B) \cap C \end{array}\right\}$ (asociatividad)
5. $\left.\begin{array}{l} A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \\ A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \end{array}\right\}$ (distributividad)

Si $\bar{X}$ representa al complementario de un conjunto X (cualquiera) respecto de un cierto conjunto universal U, dado, entonces también se cumple que:

6. $A \cup U = U \qquad A \cap U = A$
7. $A \cup \bar{A} = U \qquad A \cap \bar{A} = \emptyset$
8. $\overline{A \cup B} = \bar{A} \cap \bar{B} \qquad \overline{A \cap B} = \bar{A} \cup \bar{B}$ (leyes de Morgan)

DEMOSTRACION

Para verificar las anteriores igualdades hay que comprobar que, en cualquiera de ellas, todo elemento del conjunto del primer miembro pertenece también al conjunto del segundo miembro y recíprocamente. Veamos que así ocurre con alguna de ellas (las otras se dejan como ejercicio):

- En la segunda de las propiedades 4, los conjuntos $A \cap (B \cup C)$ y $(A \cup B) \cup (A \cap C)$ son iguales debido a que cada una de las siguientes relaciones es equivalente con la que va detrás de ella:

$$x \in A \cap (B \cup C) \quad [x \in A \text{ y } x \in B \cup C] \quad [(x \in A) \text{ y } (x \in B \text{ o } x \in C)]$$
$$[(x \in A \text{ y } x \in B) \text{ o } (x \in A \text{ y } x \in C)] \quad [x \in A \cap B \text{ o } x \in A \cap C]$$
$$x \in (A \cap B) \cup (A \cap C)$$

(*) La unión, la intersección y la complementación de conjuntos se llaman operaciones booleanas en honor al matemático George Boole (1815-1864).

- En la segunda de las propiedades 8, los conjuntos $\overline{A \cap B}$ y $\bar{A} \cup \bar{B}$ son iguales ya que, para cualquier $x \in U$, cada una de las siguientes relaciones es equivalente a la que va detrás de ella:

$$x \in \overline{A \cap B} \quad x \notin A \cap B \quad [x \notin A \text{ o } x \notin B]$$
$$[x \in \bar{A} \text{ o } x \notin \bar{B}] \quad x \in \bar{A} \cup \bar{B}$$

EJERCICIO

Compruébese que para que sea $A \cap B = \emptyset$ es necesario y suficiente que se verifique $(A \cup B) \cap \bar{B} = A$.

RESOLUCION

De acuerdo con las anteriores propiedades, es:

$$(A \cup B) \cap \bar{B} = (A \cap \bar{B}) \cup (B \cap \bar{B}) = (A \cap \bar{B}) \cup \emptyset = A \cap \bar{B}$$

Por ello, la última igualdad del enunciado se puede poner en la forma $A \cap \bar{B} = A$, que equivale a $A \subset \bar{B}$, que significa lo mismo que $A \cap B = \emptyset$.

[359]

Generalización de las operaciones booleanas. Estas operaciones se pueden definir, también, para cualquier familia de conjuntos $\{A_i / i \in I\}$ (véase [364], d): se llaman unión e intersección de los conjuntos A_i, para i recorriendo la familia de índices I, a los conjuntos:

$$\bigcup_{i \in I} A_i = \{x / \exists\, i \in I,\ x \in A_i\} \quad ; \quad \bigcap_{i \in I} A_i = \{x / \forall\, i \in I,\ x \in A_i\}$$

Estas operaciones siguen siendo asociativas y conmutativas. También se cumplen las leyes de Morgan, esto es, respecto de cualquier conjunto universal, se verifica que:

$$\complement\left(\bigcup_{i \in I} A_i\right) = \bigcap_{i \in I} (\complement A_i) \quad ; \quad \complement\left(\bigcap_{i \in I} A_i\right) = \bigcup_{i \in I} (\complement A_i)$$

APENDICE

3

Aplicaciones o funciones

Aquí, en Algebra, para caminar con soltura y seguridad, es necesario conocer con precisión lo que son las aplicaciones (o funciones). Es menester fundamentar las ideas intuitivas que, seguro, tenemos sobre estos temas. Aunque todos creamos entender lo que significa que «una función asocia a cada elemento x otro y» o que «una función es una ley que a cada x le atribuye un y», es necesario aclarar que quiere decir eso de «asociar» o lo que es esa «ley que atribuye» a cada x un y. Vamos a intentar aclarar aquí estas cuestiones y otras similares que se plantean con las funciones.

PRODUCTO CARTESIANO DE CONJUNTOS

[360]

Dados dos objetos matemáticos a y b (en este orden), se admite la existencia de un nuevo objeto que se llama *par* de primer componente a y segunda componente b y se representa poniendo (a, b); se conviene en que $(a, b) = (a', b')$ equivale a $(a = a'$ y $b = b')$. Análogamente se introducen las ternas (a, b, c), cuaternas (a, b, c, d), etc.

Se llama *producto cartesiano* de un conjunto A por otro B (en este orden) al conjunto de todos los pares (a, b) con $a \in A$ y $b \in B$, el cual se denota por $A \times B$. Análogamente se definen los productos $A \times B \times C$, $A \times B \times C \times D$, como los conjuntos que forman las correspondientes ternas, cuaternas, etc.

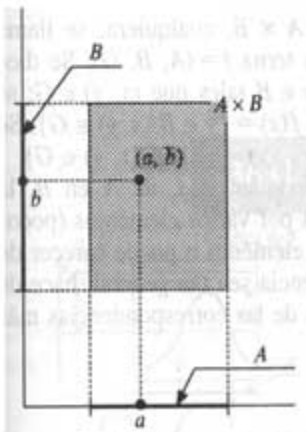


Figura 2.

OBSERVACIONES

- Un par no es otra cosa que dos elementos dados en un orden; si se cambia de elementos o si se altera el orden de éstos, se obtiene un par distinto. De manera que $(a, b) \neq (b, a)$, salvo que sea $a = b$. Un par (a, b) no se debe confundir con el conjunto $\{a, b\}$; el orden en el que figuran los elementos a y b no tiene la menor importancia en $\{a, b\}$ y, por el contrario, juega un papel fundamental en (a, b).
- En los productos cartesianos de un conjunto por sí mismo, se simplifica la notación poniendo

$$A \times A = A^2, \quad A \times A \times A = A^3, \quad \text{etc.}$$

- Al producto $A \times B$ se le califica de cartesiano por similitud con las coordenadas cartesianas(*) en el plano.

  En esta interpretación geométrica, los conjuntos A y B pasaríanm a ser (Figura 2), por ejemplo, segmentos de los ejes de abcisas y ordenadas, respectivamente; el par (a, b) con $a \in A$ y $b \in B$ vendría representado por el punto del plano cuyo sistema de coordenadas es (a, b); el producto $A \times B$ se asimilaría al paralelogramo, de lados paralelos a los ejes, que se proyecta sobre ellos según los segmentos A y B.
- Es fácil comprobar que se verifican las siguientes propiedades:

  a) Si A, A', B y B' con conjuntos no vacíos, entonces se verifica que:

$$A \times B = A' \times B' \Rightarrow (A = A' \text{ y } B = B')$$

  b) Para cualesquiera conjuntos A, A' y B, se verifica que:

$$(A \cup A') \times B = (A \cup B) \times (A' \cup B)$$
$$(A \cap A') \times B = (A \cap B) \times (A' \cap B)$$

[361]

Dados dos conjuntos A y B, se llaman *grafos* o *gráficas* relativos al producto cartesiano $A \times B$ a los subconjuntos de $A \times B$. Dada una gráfica $G \subset A \times B$, el conjunto de los $x \in A$ para los que hay algún $y \in B$ tal que $(x, y) \in G$ recibe el nombre de *proyección* de G sobre A; análogamente se define la proyección de G sobre B.

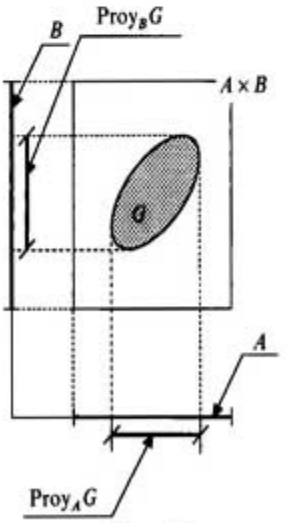


Figura 3.

OBSERVACION

Enseguida vamos a ocuparnos de las «aplicaciones» (o funciones), que son aquellas correspondencias que van a tener mayor interés para nosotros. Aunque de las correspondencias, en general, vamos a ocuparnos poco, sí debemos, al menos, dar su definición:

Dados dos conjuntos A y B y un grafo $G \subset A \times B$, cualquiera, se llama *correspondencia* de A en B, definida por G, a la terna $f = (A, B, G)$. Se dice que f hace corresponder a cada $x \in A$ todos los $y \in B$ tales que $(x, y) \in G$; se lama imagen por f del elemento $x \in A$ al conjunto $f(x) = \{y \in B / (x, y) \in G\}$. Se llama imagen recíproca de un $y \in B$ al conjunto $f^{-1}(y) = \{x \in A / (x, y) \in G\}$.

Obsérvese que, en general, dada una correspondencia f, de A en B, la imagen de un elemento $x \in A$ puede estar formada por varios elementos (pocos o muchos; incluso infinitos), puede tener un solo elemento o puede carecer de ellos. El hecho de que el concepto de correspondencia sea tan general, hace de él algo poco útil. En lo que sigue nos ocuparemos de las correspondencias más importantes: las aplicaciones.

(*) Así llamadas en honor al matemático René Descartes (1596-1650).

APLICACIONES O FUNCIONES

[362]

Sean dados dos conjuntos A y B y una gráfica $G \subset A \times B$ que goza de la siguiente propiedad: para cada elemento $x \in A$, sin excepción, existe un y sólo un elemento $y \in B$ tal que $(x, y) \in G$. Se dice entonces que G determina una *aplicación* f de A en B o una *función* definida en A y que toma valores en B (*). Para cada $x \in A$, se denota por $f(x)$ al único elemento de B para el que se verifica que $(x, f(x)) \in G$; se dice que $f(x)$ es la *imagen* de x. Se suele poner:

$$f: A \to B, \quad x \mapsto y = f(x)$$

y se simplifica la notación (si se sobrentiende quiénes son A y B) escribiendo simplemente $f: x \mapsto y = f(x)$ o sólo $x \mapsto y = f(x)$. A veces, en lugar de $f: A \to B$, se pone $A \xrightarrow{f} B$. Se dice que A es el *conjunto origen* o *campo de definición* de f y que B es su *conjunto de llegada*.

El conjunto de las aplicaciones de A en B se denota por $\mathcal{F}(A, B)$ o también por B^A.

(*) La función f se define, con precisión, como la terna $f = (A, B, G)$.

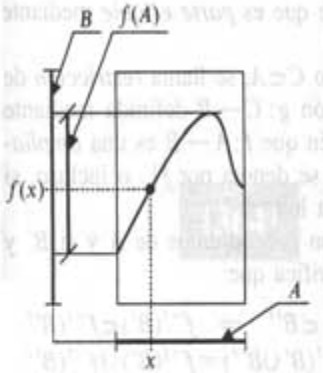


Figura 4.

COMPLEMENTOS

a) Se llama *aplicación idéntica*, para un conjunto C, a la aplicación $i: C \to C$ definida mediante $i(x) = x$ para todo $x \in C$; la gráfica de i es el conjunto $\{(x, x)/x \in C\}$, que se llama *diagonal de* $C \times C$. Si $A \subset B$, se llama *aplicación inclusión* (de A en B) a la $in: A \to B$ definida por $in(x) = x$ para todo $x \in A$.

b) Una aplicación $f: A \to B$ recibe el nombre de *aplicación constante* si todos los elementos de A tienen la misma imagen en B. Si esta imagen común fuese el elemento nulo de B (en el supuesto de que B tenga alguna cierta estructuración algebraica con elemento nulo), se dirá que f es la *aplicación nula* ($f = o$).

c) Dado un grafo cualquiera $G \subset A \times B$, se llaman *aplicaciones proyección* de G sobre A y sobre B a las:

$$Pry_A: G \to A, \quad (x, y) \mapsto x \quad \text{y} \quad Pry_B: G \to B, \quad (x, y) \mapsto y$$

[363]

Sea $f: A \to B$ una aplicación dada. Se llama imagen, por f, de un conjunto $A' \subset A$ al conjunto que forman las imágenes de los elementos de A', esto es, a $f(A') = \{f(x)/x \in A'\}$. Al conjunto $f(A) \subset B$ se le llama *imagen de la aplicación* f y se le denota por Im f; en general, Im $f \neq B$.

Se llama *imagen recíproca*, por f, de un conjunto $B' \subset B$ al conjunto que forman los elementos de A cuya imagen pertenece a B', esto es, a $f^{-1}(B') = \{x \in A/f(x) \in B'\}$. En general, f^{-1} no es una aplicación (de B en A).

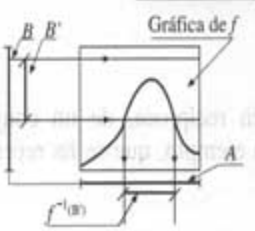


Figura 5.

OBSERVACIONES Y COMPLEMENTOS

a) Si $f: A \to B$ es una aplicación y dado un $y \in B$, se llama imagen recíproca de y por f al conjunto $f^{-1}(y) = \{x \in A / f(x) = y\}$. Cuando $f^{-1}(y)$ conste de un solo elemento $x \in A$, se considerará que es $f^{-1}(y) = x$, en lugar de $f^{-1} = \{x\}$.

b) Se llaman *transformaciones* a las aplicaciones de un conjunto en sí mismo. Dada una transformación $f: A \to A$, se dice que $x \in A$ es un *elemento fijo* de f si $f(x) = x$; un conjunto $A' \subset A$ se dice que es *parte estable* mediante f si $f(A') \subset A'$.

c) Dada una aplicación $f: A \to B$ y un conjunto $C \subset A$, se llama *restricción* de la aplicación f al conjunto C a la aplicación $g: C \to B$ definida mediante $g(x) = f(x)$ para todo $x \in C$; se dice también que $f: A \to B$ es una *ampliación* de $g: C \to B$. La restricción de f a C se denota por $f|_C$ o incluso, si ello no se presta a confusión, por la misma letra f.

d) Dada una función $f: A \to B$, si A' y A'' son subconjuntos de A y si B' y B'' son subconjuntos de B, entonces se verifica que:

1. $A' \subset A'' \Rightarrow f(A') \subset f(A'')$
2. $B' \subset B'' \Rightarrow f^{-1}(B') \subset f^{-1}(B'')$
3. $f(A' \cup A'') = f(A') \cup f(A'')$
4. $f^{-1}(B' \cup B'') = f^{-1}(B') \cup f^{-1}(B'')$
5. $f(A' \cap A'') \subset f(A') \cap f(A'')$
6. $f^{-1}(B' \cap B'') = f^{-1}(B') \cap f^{-1}(B'')$
7. $f^{-1}(f(A')) \supset A'$
8. $f(f^{-1}(B')) \subset B'$

La comprobación de estas propiedades quedará como ejercicio.

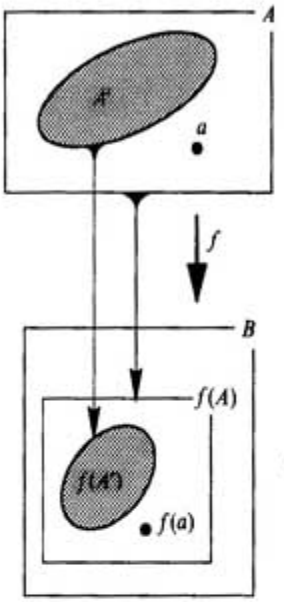


Figura 6.

EJERCICIO

Determínese el conjunto $C \subset \mathbb{R}$ más amplio para el que $f(x) = \sqrt{1 - x^2}/x^2$ define una función $f: C \to \mathbb{R}$ (esto es, hallar el campo de definición de f). Hállese tambien $\text{Im}(f) = f(C)$.

RESOLUCION

Para que $f(x)$ sea un número real, la condición es $[1 - x^2 \geq 0$ y $x^2 \neq 0]$; por tanto $C = [-1, 1] - \{0\}$. La imagen $f(C)$ es $\mathbb{R}_+ = \{y \in \mathbb{R} / y \geq 0\}$ ya que, además de ser $f(x) \geq 0$ para todo $x \in C$ y $f(1) = f(-1) = 0$, dado cualquier $y > 0$, la ecuación $y = f(x)$ tiene siempre solución para algún $x \in C$, pues equivale a

$$x = \pm \frac{(\sqrt{1 + 4y^2} - 1)^{1/2}}{\sqrt{2y^2}}$$

que tiene dos soluciones (para x) para las que evidentemente se cumple que $-1 \leq x \leq 1$ y $x \neq 0$, o sea $x \in C$.

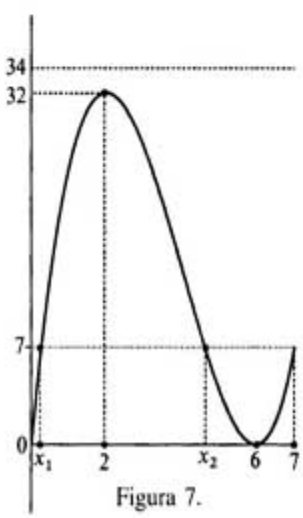


Figura 7.

EJEMPLO

Para ilustrar los conceptos de imagen e imagen recíproca, de un conjunto mediante una aplicación, considérese el siguiente ejemplo, que se ha representado en la Figura 7:

$$f: [0, 7] \to \mathbb{R}, \quad x \mapsto y = f(x) = x(x - 6)^2$$

Según se aprecia en la figura, la gráfica de f es creciente en $[0, 2[$ y en $]6, 7]$. El conjunto imagen de f es $f([0, 7]) = [0, 32]$; nótese que $f([0, 2]) = [0, 32]$, que $f([2, 6]) = [0, 32]$ y que $f([6, 7]) = [0, 7]$. La imagen recíproca de $y = 7$ está formada por las raíces de la ecuación $x(x-6)^2 = 7$, que son los valores $x_0 = 7$, $x_1 = (5 - \sqrt{21}) : 2$ y $x_3 = (5 + \sqrt{21}) : 2$. Nótese que $f^{-1}(0) = \{0, 6\}$ y que $f^{-1}(34) = \emptyset$. La imagen recíproca de $[0, 7]$ es:

$$f^{-1}([0, 7]) = [0, x_1] \cup [x_2, 7]$$

APLICACIONES SOBREYECTIVAS, INYECTIVAS Y BIYECTIVAS

Se consideran ahora aquellas aplicaciones que se comportan más regularmente:

[364]

Una aplicación $f: A \to B$ se dice que es *inyectiva* si dos elementos distintos, cualesquiera, de A tienen imágenes distintas en B; esto es, si

$$[x, x' \in A \text{ y } x \neq x'] \Rightarrow f(x) \neq f(x')$$

Una aplicación $f: A \to B$ se dice que es *sobreyectiva* si $f(A) = B$; es decir, si todo elemento de B es la imagen de algún elemento (uno al menos) de A. Una aplicación $f: A \to B$ se dice *biyectiva* si es inyectiva y sobreyectiva, o sea, si todo $y \in B$ es imagen de un y sólo un $x \in A$. Si $f: A \to B$ es biyectiva, se llama *aplicación recíproca* de f a la $f^{-1}: B \to A$ que a cada $y \in B$ le hace corresponder el único $x \in A$ tal que $f(x) = y$.

OBSERVACIONES Y COMPLEMENTOS

a) La aplicación $f: A \to B$ es inyectiva si se verifica que

$$[x, x' \in A \quad \text{y} \quad f(x) = f(x')] \Rightarrow x = x'$$

b) Para cualquiera que sea la aplicación $f: A \to B$, llamando $C = f(A)$, la aplicación $f: A \to C$ es sobreyectiva. Si $f: A \to B$ fuera inyectiva, entonces $f: A \to C$ sería biyectiva.

c) Si $f: A \to B$ es una aplicación biyectiva, entonces las relaciones $y = f(x)$ y $x = f^{-1}(y)$ son equivalentes (para $x \in A$ e $y \in B$).

d) Supóngase que $f: I \to X$ es una biyección y, para cada $i \in I$, llamemos $x_i = f(i)$. En estas condiciones, el conjunto X se puede escribir en la forma $X = \{x_i / i \in I\}$, que suele denotarse poniendo $X = \{x_i\}_{i \in I}$. Se dice entonces que I es una *familia de índices* que permite denominar un elemento cualquiera $x \in X$, recurriendo a su imagen inversa $i = f^{-1}(x)$, escribiendo $x = x_i$.

EJERCICIO

Analícese si la siguiente aplicación es sobreyectiva y si es inyectiva

$$f: \mathbb{R}^* \to \mathbb{R}, \quad f(x) = \frac{x-1}{x} \quad (\text{para } x \neq 0)$$

RESOLUCION

La aplicación no es sobreyectiva ya que, dado $y \in \mathbb{R}$, la relación $y = f(x)$ equivale a la $x = 1 : (1 - y)$, que para $y = 1$ no proporciona ningún valor de x. Nótese que $f(\mathbb{R}^*) = \mathbb{R} - \{1\}$ y que, por ello, la aplicación $f: \mathbb{R}^* \to \mathbb{R} - \{1\}$ es sobreyectiva.

La aplicación f es inyectiva, ya que

$$f(x) = f(x') \Rightarrow \frac{x-1}{x} = \frac{x'-1}{x'} \Rightarrow xx' - x' = xx' - x \Rightarrow x = x'$$

EJERCICIO

Sean dados una aplicación $f: A \to B$ y un conjunto $C \subset A$. Compruébese que:

a) Si f es sobreyectiva, entonces $f(A - C) \supset B - f(C)$.
b) Si f es inyectiva, entonces $f(A - C) \subset B - f(C)$.
c) Si f es biyectiva, entonces $f(A - C) = B - f(C)$.

RESOLUCION

a) Dado $y \in B - f(C)$, como f es sobreyectiva, ha de ser $y = f(x)$ para cierto $x \in A$ y, como $y \notin f(C)$, ha de ser $x \notin C$; en consecuencia, $y = f(x)$ para algún $x \in A - C$, esto es, $y \in f(A - C)$, como había que comprobar.
b) Si $y \in f(A - C)$, entonces existe $x \in A - C$ tal que $y = f(x)$, luego $y = f(x)$ para cierto $x \notin C$ y, como f es inyectiva, para todo $x' \in C$ ha de ser $f(x') \neq f(x) = y$, de donde resulta que $y \notin f(C)$. es decir $y \in B - f(C)$, como había que comprobar.
c) Este resultado es consecuencia obvia de los *a*) y *b*).

COMPOSICION DE APLICACIONES

[365]

Sean dos aplicaciones $f: A \to B$ y $g: C \to D$ y supóngase que $B \subset C$; en este supuesto, para cada $x \in A$, existe un único elemento $g(f(x))$ en D. Pues bien, se llama *composición* de f con g a la aplicación de A en D, que se denota poniendo $g \circ f$, definida mediante $(g \circ f)(x) = g(f(x))$; esto es:

$$g \circ f: A \to D \quad ; \quad x \mapsto (g \circ f)(x) = g(f(x))$$

La composición de aplicaciones es asociativa; no es conmutativa.

Si $f: A \to B$ es una aplicación biyectiva y llamando $i_A \to A$ e $i_B: B \to B$ a las aplicaciones idénticas, se verifica que

$$f^{-1} \circ f = i_A \quad ; \quad f \circ f^{-1} = i_B$$

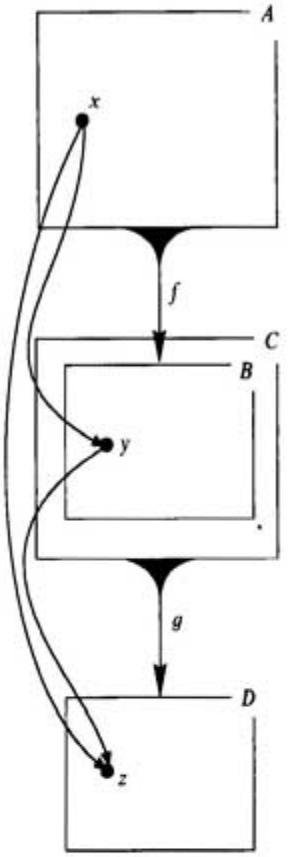


Figura 8.

OBSERVACIONES

a) La asociatividad de la composición de aplicaciones significa que, siempre que tengan sentido las composiones de aplicaciones $h\circ(g\circ f)$ y $(h\circ g)\circ f$, éstas son dos aplicaciones iguales. Así ocurre ya que, para cualquiera que sea el elemento x del conjunto origen de f, es:

$$(h\circ(g\circ f))(x) = h(g(f(x))) \quad \text{y} \quad ((h\circ g)\circ f)(x) = h(g(f(x)))$$

luego son iguales.

b) Dadas dos aplicaciones $f: A \to B$ y $g: C \to D$, para que exista $g\circ f$ es necesario que $B \subset C$ y para que exista $f\circ g$ se precisa que sea $D \subset A$; puede pues existir una de las $g\circ f$ y $f\circ g$ y no existir la otra. Aun cuando existan $g\circ f$ y $f\circ g$, en general serán distintas (la composición de aplicaciones no es conmutativa). Así, por ejemplo, si f y g son las aplicaciones, de $\mathbb{R}$ en $\mathbb{R}$, dadas por $f(x) = x^2$ y $g(x) = \text{sen } x$, será. en general es:

$$g(f(x)) = \text{sen } x^2 \neq \text{sen}^2 x = f(g(x))$$

EJERCICIO

Dadas dos aplicaciones $f: A \to B$ y $g: B \to C$, compruébese que:

a) Si f y g son inyectivas, entonces también lo es $g\circ f$.
b) Si f y g son sobreyectivas, entonces también lo es $g\circ f$.
c) Si f y g son biyectivas, entonces también lo es $g\circ f$. En este caso, compruébese que $(g\circ f)^{-1} = f^{-1}\circ g^{-1}$.

RESOLUCION

a) Si $x, x' \in A$ y es $x \neq x'$, entonces es $f(x) \neq f(x')$ (pues f es inyectiva) y de aquí se infiere (por su g inyectiva) que $g(f(x)) \neq g(f(x'))$, luego $g\circ f$ es inyectiva.

b) Por ser $f(A) = B$ y $g(B) = C$, se verifica que $(g\circ f)(A) = g(f(A)) = g(B) = C$, luego $g\circ f$ es sobreyectiva.

c) De acuerdo con *a*) y *b*), $g\circ f$ es inyectiva y sobreyectiva, luego es biyectiva, esto es, existe $(g\circ f)^{-1}$.

Para todo $z \in C$, la relación $(f^{-1}\circ g^{-1})(z) = x$ equivale a la $g^{-1}(z) = f(x)$, o lo que es lo mismo, a $z = g(f(x)) = (g\circ f)(x)$, que a su vez equivale a $(g\circ f)^{-1}(z) = x$; como para todo $z \in C$ ha resultado que $(f^{-1}\circ g^{-1})(z) = (g\circ f)^{-1}(z)$, se concluye que $f^{-1}\circ g^{-1} = (g\circ f)^{-1}$.

APENDICE

4

Relaciones de equivalencia

A no dudarlo, son muchas las relaciones de equivalencia con las que nos hemos venido ya encontrando en matemáticas. Una tal relación viene a ser algo así como un criterio que permite decir que varios objetos son equiparables o, incluso, identificables, formando todos una misma clase; con ello, en lugar de considerar elementos sueltos, se hablará de las clases que forman los elementos que sean equivalentes entre sí. Veamos bajo que condiciones las cosas son como se ha señalado.

RELACIONES BINARIAS

[366]

Dados un conjunto A y un grafo $R \subset A \times A$, se dice que un elemento $x \in A$ esta relacionado mediante R con otro $y \in A$ si $(x, y) \in R$, lo que se indica poniendo xRy; la función proposicional r dada por $r(x, y) = xRy$ se llama *relación binaria*. Dicho de otro modo: una relación binaria, entre elementos de A, es una función proposicional r, con dos variables x e y que recorren A; se llama *grafo* de r al conjunto $R = \{(x, y) \in A \times A / r(x, y)\}$. Nótese que $r(x, y)$ equivale a $(x, y) \in R$.

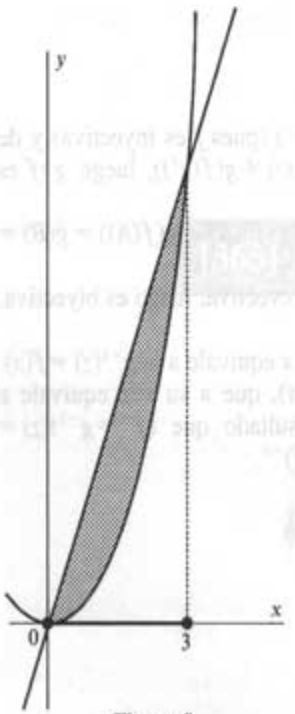


Figura 9.

OBSERVACIONES

1. Para denotar a una relación binaria se suele utilizar la misma letra con la que designa a su gráfica y se dice, por ejemplo, «sea R una relación binaria» en lugar de decir «sea r la relación binaria cuya gráfica es R».
2. Si R es una relación binaria, entre elementos de un conjunto A, un elemento $x \in A$ no estarán relacionado por R con otro elemento $y \in A$ si $(x, y) \notin R$, lo que se expresa poniendo $x\bar{R}y$, y equivale a decir que $x\bar{R}y$, donde $\bar{R} = A \times A - R$ (complementario de R).

EJEMPLO

Consideremos la relación R definida, entre números reales, mediante

$$xRy \Leftrightarrow x^2 \leqslant y \leqslant 3x$$

La gráfica de esta relación es el conjunto de los puntos $(x, y) \in \mathbb{R}^2$ que están

ubicados en la región limitado entre la parábola $y = x^2$ y la recta $y = 3x$ (para $0 \leqslant x \leqslant 3$) incluyendo la periferia (véase la Figura). Los elementos que están relacionados consigo mismos son los del intervalo [0, 1] y sólo ellos. En esta relación no hay simetría, esto es, xRy no equivale a yRx. Así ocurre ya que, tomando, por ejemplo, $x = 2$ e $y = 5$, se verifica que $2R5$ pero $5\not R2$.

EJERCICIO

Considérese la relación R, entre números reales no nulos (esto es, definida en $\mathbb{R}^*$), dada mediante:

$$[xRy] \Leftrightarrow [\text{los vectores } \bar{u} = (x, y) \text{ y } \bar{v} = (y, x) \text{ forman un ángulo de } 30°]$$

a) Obtener la gráfica de R.
b) Hallar las posibles $a \in \mathbb{R}^*$ tales que aRa.
c) Probar que dado $a \neq 0$ existe algún $b \neq 0$ tal que aRb y bRa.

RESOLUCION

a) Recurriendo a que

$$\cos(\bar{u}, \bar{v}) = \frac{xy + yx}{\sqrt{x^2 + y^2}\sqrt{y^2 + x^2}}$$

la relación R se puede expresar

$$[xRy] \Leftrightarrow 2xy = \frac{\sqrt{3}}{2}(x^2 + y^2) \Leftrightarrow (\sqrt{3}x - y)(x - \sqrt{3}y) = 0 \Leftrightarrow$$

$$\Leftrightarrow [y = \sqrt{3}x \quad \text{o} \quad y = x/\sqrt{3}]$$

La gráfica de R está, pues constituida por las rectas $y = \sqrt{3}x$ e $y = x/\sqrt{3}$ (excepto el punto $x = 0$, $y = 0$).

b) No hay ningún elemento $a \neq 0$ tal que aRa, ya que $a = \sqrt{3}a$ o $a = a/\sqrt{3}$ sólo se verifican para $a = 0$.

c) Dado $a \neq 0$, las relaciones aRb y bRa se verifican, ambas, para $b = a\sqrt{3}$ y para $b = a/\sqrt{3}$.

[367]

Sea R una relación binaria entre elementos de un conjunto A. Se dice que R es reflexiva, que es simétrica, que es antisimétrica o que es transitiva si se verifica, respectivamente:

reflexiva:	xRx,	$\forall x \in A$
simétrica:	$xRy \Rightarrow yRx$,	$\forall x, y \in A$
antisimétrica:	$(xRy, yRx) \Rightarrow x = y$,	$\forall x, y \in A$
transitiva:	$(xRy, yRz) \Rightarrow xRz$,	$\forall x, y, z \in A$

EJEMPLOS

1. Entre los números enteros, vamos a considerar la relación binaria R definida mediante

$$mRn \Leftrightarrow m+n \text{ es múltiplo de } 6$$

Esta relación no es reflexiva ya que hay valores de m para los que $m\not Rm$ (por ejemplo, $m=2$), aunque para otros sea mRm (por ejemplo, $m=9$). Esta relación es simétrica ya que mRn implica nRm. Esta relación no es antisimétrica ya que (tomando $m=2$ y $n=8$) es $2R8$ y $8R2$ aun cuando $2 \neq 8$. La relación tampoco es transitiva, ya que (por ejemplo) es $7R5$ y $5R1$ pero $7\not R1$.

2. Sea ahora la relación definida, entre números reales, como sigue:

$$xRy \Leftrightarrow x \leqslant y < x+3$$

Aquí, R es reflexiva ya que $x \leqslant x < x+3$ para todo $x \in \mathbb{R}$. Esta relación no es simétrica puesto que, por ejemplo, $5R6$ para $6\not R5$. La relación es antisimétrica, puesto que, de ser $x \leqslant y < x+3$ e $y \leqslant x < y+3$, deberá verificarse que $x=y$. Finalmente, la relación no es transitiva ya que, por ejemplo, es $4R6$ y $6R8$ pero $4\not R8$.

OBSERVACIONES

Según se puede comprobar con facilidad, se verifica que:

1. Para que una relación binaria R, definida entre los elementos de un conjunto A, sea simétrica y antisimétrica es necesario y suficiente que para cualquier $x \in A$ sólo quepa una de las dos posibilidades siguientes: 1) x está relacionado solamente consigo mismo; y 2) x no está relacionado con ningún elemento de A.
2. Si R es una relación binaria, entre elementos de un conjunto A, que es reflexiva y transitiva, entonces la nueva relación S definida mediante

$$(xSy) \Leftrightarrow [xRy \text{ e } yRx]$$

es reflexiva, simétrica y transitiva.

3. Sea S una relación binaria definida en un conjunto B y sea $f: A \to B$ una aplicación. Se define en A la relación binaria R dada por

$$[xRy] \Leftrightarrow [f(x)Sf(y)]$$

Si S es reflexiva, simétrica o transitiva, entonces también lo es R. Si S es antisimétrica y f es inyectiva, entonces R también es antisimétrica.

RELACIONES DE EQUIVALENCIA

[368]

Una relación binaria, entre elementos de un conjunto A, que sea reflexiva, simétrica y transitiva recibe el nombre de *relación de equivalencia* o, simplemente, equivalencia. Generalmente una equivalencia se denota por el signo $\sim$; para indicar que x está relacionado con y se pone «$x \sim y$» y se lee «x es equivalente a y».
El conjunto de los elementos de A que son equivalentes a un $x \in A$ se denota por

$$[x] = \{x' \in A / x' \sim x\}$$

y recibe el nombre de *clase de equivalencia* de representante x.

EJEMPLOS

1. Para el conjunto $\mathbb{R}^2$, la relación $\sim$ definida mediante

$$(x, y) \sim (x', y') \quad \Leftrightarrow \quad xy = x'y'$$

es una equivalencia ya que para cualesquiera (x, y), (x', y') y (x'', y'') de $\mathbb{R}^2$ se verifica que:

- $xy = xy$, luego la relación $\sim$ es reflexiva.
- $xy = x'y' \Rightarrow x'y' = xy$, luego $\sim$ es simétrica.
- $(xy = x'y' \text{ y } x'y' = x''y'') \Rightarrow xy = x''y''$, luego $\sim$ es transitiva.

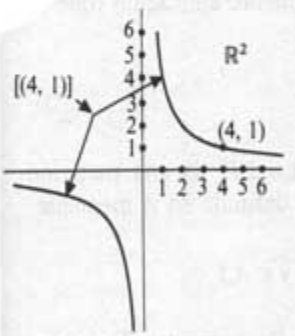


Figura 10.

Dado un $(x_0, y_0) \in \mathbb{R}^2$ cualquiera, la clase de equivalencia que tiene por representante a (x_0, y_0) está formada por todos los puntos (x, y) de $\mathbb{R}^2$ (véase la Figura 9) tales que $xy = k$, donde k es el número real $k = x_0 y_0$. Esta clase de equivalencia es, en la interpretación geométrica usual de $\mathbb{R}^2$, una hipérbola equilátera.

2. Dado un número natural $p > 1$, consideremos la relación binaria $\sim$ definida entre los números enteros mediante:

$$[m \sim n] \quad \Leftrightarrow \quad [m - n \text{ es múltiplo de } p]$$

(en lugar de $m \sim n$, en este caso se suele poner $m \equiv n(p)$ y se dice que m y n son «congruentes modulo p»).

Esta relación es una equivalencia, ya que:

- Es reflexiva, pues $n - n = 0$ es múltiplo de p para todo $n \in \mathbb{Z}$.
- Es simétrica, puesto que, si $m - n$ es múltiplo de p, entonces $n - m = -(m - n)$ también es múltiplo de p.
- Es transitiva, ya que, si $m - n$ y $n - h$ son múltiplos de p, entonces $m - h = (m - n) + (n - h)$ también es múltiplo de p.

En esta equivalencia, la clase a la que pertenece un elemento $m \in \mathbb{Z}$, que la denotaremos poniendo $\overline{m}$, está formada por todos los enteros que, al dividirlos entre p, conducen al mismo resto que m. Como el resto, de dividir entre p, tiene que ser uno de los $0, 1, 2, ..., p-1$, resulta que hay p clases de equivalencia, que son las $\overline{0}, \overline{1}, \overline{2}, ..., \overline{p-1}$. Nótese que los enteros que forman la clase $\overline{r}$ son los que pueden ponerse en la forma $\dot{p} + r$, donde $\dot{p}$ denota a un múltiplo entero de p.

CONJUNTO COCIENTE

[369]

Teorema.—Si $\sim$ es una equivalencia entre elementos de un conjunto A, entonces para cualesquiera $x, y \in A$ se verifica que

$$(x \sim y) \Leftrightarrow [x] = [y] \qquad no(x \sim y) \Leftrightarrow [x] \cap [y] = \emptyset$$

Las clases de equivalencia forman una partición propia(*) de A. La familia de las clases de equivalencia recibe el nombre de *conjunto cociente* y se denota poniendo $A/\sim$; esto es:

$$A/\sim = \{[x]/x \in A\} \quad \text{donde} \quad [x] = \{x' \in A / x' \sim x\}$$

Se llama *aplicación canónica* o natural a la siguiente aplicación (que es sobreyectiva), de A en el conjunto cociente $A/\sim$:

$$A \to A/\sim \quad x \mapsto [x]$$

Recíprocamente, si la familia de conjuntos $\{A_i / i \in I\}$ es una partición propia(*) del conjunto A, entonces la relación $\sim$ definida en A mediante

$$(x \sim y) \Leftrightarrow (\exists\, i \in I,\ x \in A_i \text{ e } y \in A_i)$$

es una equivalencia en A que tiene por clases a los conjuntos A_i dados.

(*) Partición propia de A es toda familia de subconjuntos no vacíos de A, que son disjuntos dos a dos y tales que su unión es A.

DEMOSTRACION

- Si es $x \sim y$, entonces para cualquier $x' \in [x]$ es $x' \sim x$ y, por la transitividad de $\sim$, sería entonces $x' \sim y$, esto es, ha de ser $x' \in [y]$, luego $[x] \subset [y]$; de modo análogo se comprueba que $[y] \subset [x]$ y, por tanto, es $[x] = [y]$.
- Recíprocamente, si $[x] = [y]$, entonces $x \in [y]$, o sea, ha de ser $x \sim y$.
- En el caso de que $no(x \sim y)$, si existiese algún $z \in [x] \cap [y]$, habría de ser $z \sim x$ y $z \sim y$ y entonces, por la simetría y por la transitividad de $\sim$, debería ser

$x \sim y$, que es falso y, por ello, $[x] \cap [y]$ debe carecer de elementos, o sea, $[x] \cap [y] = \emptyset$.

- Recíprocamente, de ser $[x] \cap [y] = \emptyset$, sería $x \notin [y]$, o sea $no(x \sim y)$.

En relación con las afirmaciones que, además de las contenidas en el teorema, se hacen en el enunciado, debemos observar que:

1. El conjunto cociente es una partición propia de A ya que las clases de equivalencia son subconjuntos no vacíos de A, disjuntos dos a dos y tales que su unión es A, pues todo $x \in A$ pertenece a la clase $[x]$.
2. Si $\{A_i / i \in I\}$ es una partición propia de A, entonces la relación $\sim$:

$$(x \sim y) \quad \Leftrightarrow \quad (x \text{ e } y \text{ pertenecen a un mismo } A_i)$$

es evidentemente reflexiva, simétrica y transitiva, esto es, se trata de una equivalencia.

EJEMPLOS

1. Para el conjunto de las rectas del plano, el paralelismo es una relación de equivalencia. Una clase de equivalencia, formada por todas las rectas que son paralelas a una recta dada, se llama la *dirección* de dicha recta.
2. Sea A el conjunto de los vectores fijos (segmentos orientados del espacio geométrico ordinario). La equipotencia de vectores (es decir, la relación binaria $\sim$ según la cual $\overrightarrow{AB} \sim \overrightarrow{CD}$ equivale a que $ABDC$ sea un paralelogramo) es una relación de equivalencia en A; las clases de esta equivalencia se llaman *vectores libres*. Un vector libre es, pues, el conjunto de todos los vectores fijos equivalentes a uno dado.
3. En el conjunto $\mathbb{N}^2$, la relación binaria $\sim$ definida por

$$(n, m) \sim (n', m') \quad \Leftrightarrow \quad n + m' = m + n'$$

es una relación de equivalencia cuyas clases son:

Si $n > m$,

$$[(n, m)] = [(d, 0)] = \{(d, 0), (d+1, 1), (d+2, 2), \ldots\} \quad \text{con } d = n - m$$

Si $n < m$,

$$[(n, m)] = [(0, d)] = \{(0, d), (1, d+1), (2, d+2), \ldots\} \quad \text{con } d = m - n$$

Si $n = m$,

$$[(n, m)] = [(0, 0)] = \{(0, 0), (1, 1), (2, 2), \ldots\}$$

El conjunto cociente está formado por todas estas clases:

$$\mathbb{N}^2/\sim = \{[(0, 0)], [(1, 0)], [(0, 1)], \ldots, [(d, 0)], [(0, d)], \ldots\}$$

Como la clase $[(d, 0)]$ se identifica con el número natural d y la clase

$[(0, d)]$ es el número entero $-d$, resulta que el anterior conjunto cociente es el conjunto $\mathbb{Z} = \{0, 1, -1, \ldots, d, -d, \ldots\}$ de los *números enteros*.

DESCOMPOSICION CANONICA DE UNA APLICACION

[370]

Sean dados dos conjuntos no vacíos A y B y una aplicación cualquiera $f: A \to B$. La relación binaria $\sim$ definida en A mediante

$$x \sim x' \Leftrightarrow f(x) = f(x')$$

es una equivalencia, que se llama *equivalencia asociada a la aplicación f*. El conjunto cociente, que se denota poniendo A/f (en lugar de $A/\sim$), es:

$$A/f = \{[x]/x \in A\}; \quad \text{donde} \quad [x] = \{x' \in A / f(x) = f(x')\}$$

Recurriendo a las siguientes aplicaciones $\tilde{f}$, i (inclusión) y c (canónica) definidas mediante:

$\tilde{f}: A/f \to f(A)$	$i: f(A) \to B$	$c: A \to A/f$
$[x] \mapsto f(x)$	$y \mapsto y$	$x \mapsto [x]$
(biyectiva)	(inyectiva)	(sobreyectiva)

se verifica que la aplicación dada f se puede descomponer en la forma:

$$f = i \circ \tilde{f} \circ c$$

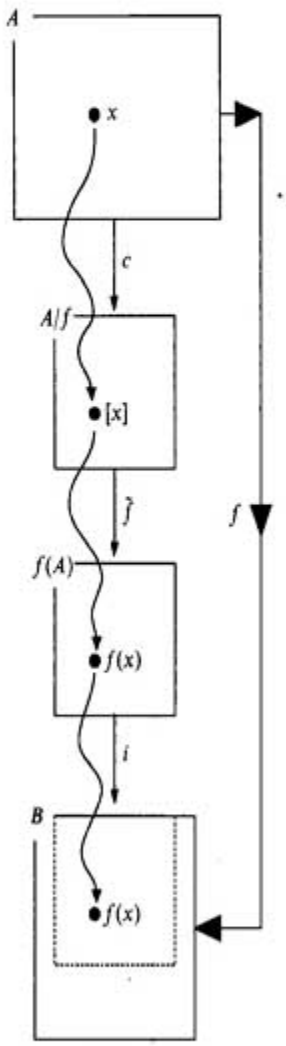


Figura 11.

DEMOSTRACION

La comprobación de que la relación $\sim$ es una equivalencia no tiene ninguna dificultad; queda como ejercicio. Nótese que la aplicación $\tilde{f}$ está definida sin ambigüedad, esto es, $\tilde{f}([x])$ no depende del representante x que se tome para definir la clase $[x]$; ello es así, porque si $[x] = [x']$, entonces $f(x) = f(x')$. Es también fácil comprobar que $\tilde{f}$ es biyectiva y, por ello, se omite. Finalmente, para comprobar que $f = i \circ \tilde{f} \circ c$, ya que: tomando un $x \in A$ cualquiera, se tiene:

$$(i \circ \tilde{f} \circ c)(x) = (i \circ \tilde{f})(c(x)) = (i \circ \tilde{f})([x]) = i(\tilde{f}([x])) = i(f(x)) = f(x)$$

EJEMPLO

Considérese la aplicación

$$f: \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y) = x^2 + y^2$$

La equivalencia asociada a f es la $\sim$ definida, entre elementos de $\mathbb{R}^2$ por:

$$(x, y) \sim (x', y') \Leftrightarrow x^2 + y^2 = x'^2 + y'^2$$

El correspondiente conjunto cociente $\mathbb{R}^2/f$ está formado por las circunferencias (de $\mathbb{R}^2$) que tienen su centro en el origen; esto es:

$$\mathbb{R}^2/f = \{C_r / r \geq 0,\ r \in \mathbb{R}\} \quad \text{donde} \quad C_r = \{(x, y) \in \mathbb{R}^2 / x^2 + y^2 = r^2\}$$

La aplicación canónica es la

$$C: \mathbb{R}^2 \to \mathbb{R}^2/f, \quad (x, y) \mapsto c(x, y) = C_r \quad \text{donde } r = \sqrt{x^2 + y^2}$$

La aplicación biyectiva $\tilde{f}$ asociada a f es aquí:

$$\tilde{f}: \mathbb{R}/f \to f(\mathbb{R}^2), \quad C_r \mapsto \tilde{f}(C_r) = r$$

APENDICE

5

Relaciones de orden

La relación «menor o igual» entre números naturales es el modelo que, debidamente generalizado, conduce a la noción de relación de orden. Esta generalización se hace de modo que se respeten las propiedades específicas de la referida relación; con tal de que así sea, poco importa el modo concreto de definirla.

RELACION DE ORDEN

[371]

Se dice que una relación binaria (véase [368]), definida en un conjunto A y que denotaremos recurriendo al signo $\leqslant$, es una *ordenación* o una relación de orden si es reflexiva, antisimétrica y transitiva; en tal caso, se dice que $(A, \leqslant)$ es un *conjunto ordenado*. La proposición «$x \leqslant y$» se lee «x es menor o igual que y» o «x es anterior a y».
Si $(A, \leqslant)$ es un conjunto ordenado, se dirá que $\leqslant$ es un *orden total* si, para cualesquiera $x, y \in A$, se verifica que $x \leqslant y$ o que $x \geqslant y$ (esto es, que x e y son comparables); si $\leqslant$ no es un orden total, se dice que es un *orden parcial*.

EJEMPLOS

1. Las ordenaciones usuales de los campos numéricos $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ y $\mathbb{R}$ son relaciones de orden totales; en ellas se inspira la anterior definición de conjunto ordenado.
2. Para el conjunto $\mathcal{P}(U)$, de las partes de un conjunto U, la relación $\leqslant$ dada por $A \leqslant B \Leftrightarrow A \subset B$ (relación de inclusión) es una ordenación; esta relación de orden es pacial (salvo si U es el vacío o tiene un solo elemento).
3. Para el conjunto $A = \mathcal{F}(C, \mathbb{R})$, de las funciones definidas en un conjunto C y que toman valores reales, considérese la relación $\leqslant$ siguiente:

$$(f \leqslant g) \quad \Leftrightarrow \quad [\forall x \in C, f(x) \leqslant g(x)]$$

Esta relación es un orden para el conjunto A y dicho orden lo es parcial (siempre que C tenga más de un elemento).

OBSERVACIONES

- Si $(A, \leqslant)$ es un conjunto ordenado, la relación ($x \leqslant y$ y $x \neq y$), para $x, y \in A$, se lee «x es estrictamente menor que y».

- Si $\leqslant$ es una relación de orden para un conjunto A, entonces también lo es para cualquier subconjunto de A.
- Si $(A, \leqslant)$ es un conjunto ordenado y dados dos elementos $a, b \in A$ tales que $a \leqslant b$, se llaman *intervalos* de origen a y extremo b a los siguientes conjuntos:

$$[a, b] = \{x \in A / a \leqslant x \leqslant b\} \quad]a, b[= \{x \in A / a < x < b\}$$
$$[a, b[= \{x \in A / a \leqslant x < b\} \quad]a, b] = \{x \in A / a < x \leqslant b\}$$

- Si $(A, \leqslant)$ es un conjunto ordenado, se dice que $C \subset A$ es una *cadena* si $(C, \leqslant)$ es un conjunto totalmente ordenado. Así, para el conjunto $\mathcal{F}(I, \mathbb{R})$ de las funciones reales definidas en el intervalo real [0, 1] y con la ordenación del ejemplo 3, las funciones $x \mapsto f_n(x) = x^n$ para $n = 1, 2, 3, \ldots$ forman una cadena.

APLICACIONES MONOTONAS

[372]

Sean dados dos conjuntos ordenados $(A, \leqslant)$ y $(B, \leqslant)$ y una aplicación $f: A \to B$. Se dice que f es una aplicación monótona si es creciente o si es decreciente, conforme a las siguientes definiciones:

f es *creciente* si: $(x, y \in A, x < y) \Rightarrow f(x) \leqslant f(y)$

f es *decreciente* si: $(x, y \in A, x < y) \Rightarrow f(x) \geqslant f(y)$

Si los anteriores signos $\leqslant$ y $\geqslant$ se sustituyen por $<$ y $>$, respectivamente, la monotonía de f se dice que es estricta.

Una aplicación entre conjuntos ordenados se dice que es un *isomorfismo respecto del orden* si es biyectiva y monótona; dos conjuntos ordenados se dicen isomorfos si entre ellos existe un isomorfismo.

EJEMPLOS

1. Consideremos la función «parte entera», esto es, la $e: \mathbb{R} \to \mathbb{Z}$ que a cada $x \in \mathbb{R}$ le asigna su parte entera $e(x)$. Suponiendo que en $\mathbb{R}$ y en $\mathbb{Z}$ se dispone de la relación de orden usual, la aplicación e es creciente, pues $x < y \Rightarrow e(x) \leqslant e(y)$, pero no es estrictamente creciente.
2. La aplicación $x \mapsto \operatorname{arctg} x$, definida en $\mathbb{R}$ y que toma sus valores en el intervalo real $I =]-\pi/2, \pi/2[$, es un isomorfismo respecto del orden, de $\mathbb{R}$ en I.

CONJUNTOS ACOTADOS

[373]

Sea $(A, \leqslant)$ un conjunto ordenado y considérese un conjunto no vacío $C \subset A$. Se dice que un elemento $k \in A$ es una *cota superior* de C si $k \geqslant x$ para todo x $\in$ C; se dice que $k \in A$ es una *cota inferior* de C si $k \leqslant x$ para todo $x \in C$. El conjunto C se dice *acotado* superior o inferiormente si tiene alguna cota superior o inferior, respectivamente.

OBSERVACIONES

- En general, las cotas de un conjunto (si las hay) podrán ser elementos que no pertenezcan al conjunto. Si k es una cota superior (inferior) de un conjunto C, también lo son todos los elementos de A mayores (menores) que k.
- Dados un conjunto ordenado $(A, \leqslant)$ y un subconjunto $C \subset A$, se dice que un elemento $c \in C$ es un elemento *maximal* (*minimal*) de C si en C no hay ningún elemento que sea mayor (menor) que c. Así, por ejemplo, para el conjunto A de los intervalos reales ordenados por inclusión, el conjunto $C = \{I, J, K\}$, donde

$$I = [-2, 5] \quad , \quad J = [0, 7] \quad \text{y} \quad K = [1, 4]$$

tiene como elementos maximales a los I y J; nótese que ninguno de ellos es cota superior de C. El intervalo K es elemento minimal de C; este intervalo es, al tiempo, cota inferior de C.

EJEMPLO

En $\mathbb{R}$ y con la relación de orden usual, considérese el conjunto C formado por los números

$$\frac{3n^2 + 6nm + 5m^2 + 10}{2n^2 + 4nm + m^2 + 6}, \quad \text{para } n, m \in \mathbb{N}$$

El conjunto C está acotado inferiormente, ya que todos sus elementos son, obviamente, positivos y también esté acotado superiormente, ya que:

$$\frac{3n^2 + 6nm + 5m^2 + 10}{2n^2 + 4nm + m^2 + 6} < \frac{(3n^2 + 6nm + 5m^2 + 10) + (7n^2 + 14nm + 20)}{2n^2 + 4nm + m^2 + 6} = 5$$

Resulta, pues, que 0 es una cota inferior de C y que 5 es una cota superior.

EXTREMOS

[374]

Sea $(A, \leqslant)$ un conjunto ordenado y considérese un subconjunto $C \subset A$.
Si C está acotado superiormente y, además, existe la menor de todas sus cotas superiores, a ella se la llama extremo superior o *supremo* de C y se la denota por sup C. Cuando existe este supremo y pertenezca a C, se le llamará *máximo* de C y se denotará poniendo máx C.
Si C está acotado inferiormente y, además, existe la mayor de todas sus cotas inferiores, a ella se la llama extremo inferior o *ínfimo* de C y se la denota por inf. C. Cuando exista este ínfimo y pertenezca a C, se le llamará *mínimo* de C y se denotará poniendo mín C.

OBSERVACION

- Si $(A, \leqslant)$ es un conjunto ordenado y dado $C \subset A$, conviene notar que el máximo y el mínimo de C son, caso de existir, los siguientes elementos. El máximo de C es, si le hay, un elemento de C que es mayor que el resto de los elementos de C. El mínimo de C es, si le hay, un elemento de C que es menor que el resto de los elementos de C.

EJEMPLOS

1. En $\mathbb{R}$ y con la ordenación usual, consideremos el conjunto $C = [0, 1[$. Este conjunto está acotado superior e inferiormente (2 es una cota superior y -5 es una cota inferior). Es evidente que sup $C = 1$ y que inf $C = 0$. Como 0 es de C, resulta que mín $C = 0$; como 1 no es de C, resulta que no existe el máximo de C.
2. Consideremos ahora el conjunto $A = \mathcal{F}([0, 1], \mathbb{R})$, de las funciones reales definidas en el intervalo real $[0, 1]$, y que, en él, se dispone de la relación de orden $\leqslant$ dada por:

$$(f \leqslant g) \Leftrightarrow (f(x) \leqslant g(x), \forall x \in [0, 1])$$

Sea C el subconjunto de A formado por las siguientes funciones f_n para $n \in \mathbb{N}$:

$$f_1(x) = x, \quad f_2(x) = x^2, \quad \ldots, \quad f_n(x) = x^4, \ \ldots$$

El conjunto C está acotado superiormente y tiene supremo, que es su máximo: máx C = sup $C = f_1$. El conjunto C está acotado inferiormente y admite como ínfimo a la función f dada por

$$f(x) = 0 \quad \text{para} \quad x \neq 1 \quad , \quad f(1) = 1$$

como esta función no es de C, este conjunto carece de mínimo.

FUNCIONES ACOTADAS

[375]

Si $(B, \leqslant)$ es un conjunto ordenado, se dirá que una función $f: A \to B$ es una *función acotada* superior o inferiormente si su conjunto imagen $f(A)$ está, respectivamente, acotado superior o inferiormente (en B); si existen sup $f(A)$ o inf $f(A)$, a ellos se les llamará supremo e ínfimo de f, respectivamente. Se llaman, en el caso de que existan, máximo y mínimo de f a los máx $f(A)$ y mín $f(A)$; esto es, se dice que m = máx f o que m = mín f si existe algún $a_0 \in A$ tal que $m = f(a_0)$ y $m \geqslant f(a)$ o $m \leqslant f(a)$, respectivamente, para todo $a \in A$.

APENDICE

6

Cardinal de un conjunto

Aunque sea poco, algo hemos de decir acerca de la clasificación de los conjuntos en función de su «número de elementos» o, mejor dicho, de su cardinal: dos conjuntos tienen el mismo cardinal si se pueden poner en correspondencia biyectiva. Además de los conjuntos finitos, que son aquellos cuyo cardinal es un número natural, nos vamos a encontrar con los conjuntos infinitos; éstos se van a caracterizar por el hecho, aparentemente contradictorio, de que se pueden poner en correspondencia biyectiva con alguna de sus partes propias. De los conjuntos infinitos, sólo nos van a interesar los llamados conjuntos numerables (que tienen el mismo cardinal que $\mathbb{N}$) y los conjuntos con la potencia del continuo (que son los que tienen igual cardinal que $\mathbb{R}$).

CONJUNTOS EQUIPOTENTES

[376]

Se dice que dos conjuntos A y B tienen igual *potencia*, que son equipotentes o que son *conjuntos coordinables* si existe una biyección entre ellos, lo que se expresa poniendo $A \simeq B$. Esta relación $\simeq$ es reflexiva, simétrica y transitiva; por ello, los conjuntos se clasifican respecto de su potencia en clases disjuntas. Dado un conjunto A, a él y a cualquiera de los conjuntos equipotentes con él se le asignará un objeto matemático llamado *cardinal* de A (Card A); dos conjuntos son equipotentes si y sólo si tienen el mismo cardinal, esto es, Card A = Card B equivale a poner $A \simeq B$.

OBSERVACIONES

1. Antes que nada, nótese que la relación $\simeq$ de equipotencia es efectivamente: 1) reflexiva, pues para todo conjunto A es $A \simeq A$ ya que la aplicación idéntica es una biyección de A en A; 2) simétrica, esto es, $A \simeq B \Rightarrow B \simeq A$, ya que si $f: A \to B$ es una biyección entonces también lo es $f^{-1}: B \to A$; y 3) transitiva, esto es, si $A \simeq B$ y $B \simeq C$ entonces $A \simeq C$, ya que si $f: A \to B$ y $g: B \to C$ son biyecciones, también lo es $g \circ f: A \to C$. La equipotencia entre conjuntos satisface, pues, a las tres leyes de las relaciones de equivalencia.
2. Para cualquiera que sea la familia $\mathcal{F}$ de conjuntos, la equipotencia es una relación de equivalencia en $\mathcal{F}$ y la clase de equivalencia de representante $A \in \mathcal{F}$ será:

$$[A] = \{B \in \mathcal{F} / \text{Card } B = \text{Card } A\}$$

Si para $\mathcal{F}$ se pudiera tomar la «familia de todos los conjuntos» (véase [421], NOTA 5, *a*), se podría haber definido el cardinal como Card $A = [A]$. Al no existir el conjunto (o familia) de todos los conjuntos, para definir los cardinales no se ha podido proceder de dicha manera y se ha hablado del cardinal como un cierto objeto matemático que comparten todos los conjuntos equipotentes a uno dado. Para salvar esta imprecisión, podría considerarse que el cardinal de un conjunto A es uno cualquiera de los conjuntos equipotentes con A.

CONJUNTOS FINITOS

[377]

A partir del conjunto vacío, construyamos los conjuntos:

$$\emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\},\ \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$$

donde los elementos de cada conjunto son los del conjunto anterior y, además, el propio conjunto anterior. Los cardinales de estos conjuntos se representan por los símbolos 0, 1, 2, 3, ..., respectivamente, y son los *números naturales*; el conjunto de todos ellos se denota por $\mathbb{N}$.
A los conjuntos que tienen por cardinal a un número natural se les llama *conjuntos finitos*.

OBSERVACIONES

1. Si A es un conjunto finito, a su cardinal se le conoce también como el *«número de elementos»* de A; dos conjuntos finitos son, pues, equipotentes si y sólo si tienen el mismo número de elementos.
2. Si A es un conjunto finito y B está estrictamente incluido en A, entonces no existe ninguna biyección entre A y B; esto es, un conjunto finito no es equipotente a ninguna de sus partes propias. Esta propiedad, a pesar de parecer evidente, no es fácil de demostrar (véase [425], nota 9, a).
3. Según lo anterior, si un conjunto A se puede poner en correspondencia biyectiva con alguna de sus partes propias, entonces A no puede ser finito; se dirá entonces que A es infinito.

EJEMPLOS

1. El conjunto $\mathbb{N}$ de los números naturales es infinito (no finito) pues es equipotente con algún subconjunto estrictamente incluido en $\mathbb{N}$. Comprobemos que así es, estableciendo una biyección entre $\mathbb{N}$ y el conjunto $P = \{2n/n \in \mathbb{N}\}$ de los números pares, que está contenido estrictamente en $\mathbb{N}$; una tal biyección puede ser, por ejemplo, la siguiente:

$$f: \mathbb{N} \to P, \quad n \mapsto f(n) = 2n$$

2. El conjunto $\mathbb{R}$ de los números reales es infinito (no finito) pues es equipotente a alguna de sus partes propias. Así ocurre, pues $\mathbb{R}$ se puede poner en correspondencia biyectiva con el intervalo $]-1, 1[$, por ejemplo, que está estrictamente incluido en $\mathbb{R}$, ya que la aplicación

$$f:]-1, 1[\to \mathbb{R}, \quad x \mapsto f(x) = \frac{x}{1-x^2}$$

es una biyección, como se comprueba fácilmente.

CONJUNTOS INFINITOS

[378]

Se llaman *conjuntos infinitos* a los conjuntos que no son finitos; es decir, un conjunto A es infinito si Card A no es un número natural.

Caracterizaciones de los conjuntos infinitos: un conjunto A es infinito si y sólo si se verifica una cualquiera de las condiciones:

- Algún subconjunto de A es equipotente con $\mathbb{N}$.
- El conjunto A es equipotente con alguna de sus partes propias.

OBSERVACIONES

1. Las anteriores caracterizaciones se comprueban aparte (véase [425], NOTA 9. *b*).
2. Los anteriores ejemplos 1 y 2 ponen de manifiesto que $\mathbb{N}$ y $\mathbb{R}$ son conjuntos infinitos. Los cardinales de estos conjuntos son, de entre los no finitos, los de mayor interés; nosotros no vamos a precisar de ningún otro cardinal infinito.

POTENCIAS DE LO NUMERABLE Y DEL CONTINUO

[379]

El conjunto $\mathbb{N}$ de los números naturales es infinito; a su cardinal se le denota por $\aleph_0$ (alef cero). Se dice que A es un *conjunto infinito numerable* si es equipotente con $\mathbb{N}$, es decir, si Card $A = \aleph_0$. Los conjuntos $\mathbb{Z}$ y $\mathbb{Q}$ son infinitos numerables.

El conjunto $\mathbb{R}$ de los números reales es infinito; $\mathbb{R}$ no es infinito numerable. Se dice que $\mathbb{R}$ y todos los conjuntos equipotentes con $\mathbb{R}$ tienen la *potencia del continuo*; el cardinal de $\mathbb{R}$ se denota por $\aleph_1$ (alef uno) o por $\mathfrak{c}$ (c gótica). El conjunto $\mathbb{C}$ tienen la potencia del continuo.

OBSERVACION

De acuerdo con la anterior definición, si A es un conjunto infinito numerable, entonces existe alguna biyección

$$f:\mathbb{N}\to A, \quad n\mapsto f(n)$$

Llamando $a_n = f(n)$, el conjunto A se puede poner:

$$A = \{a_n / n \in \mathbb{N}\} = \{a_0, a_1, a_2, ..., a_n, ...\}$$

Recíprocamente, todo conjunto de la forma $A = \{a_0, a_1, a_2, ..., a_n, ...\}$ (con $a_i \neq a_j$ para $i \neq j$) es infinito numerable ya que $f:\mathbb{N}\to A$, $f(n) = a_n$, es una biyección.

Podemos resumir lo anterior diciendo que un conjunto A es infinito numerable si y sólo si es posible numerar sus elementos recurriendo a todos los números naturales o, también, si y sólo sí los elementos de A pueden ponerse en forma de *sucesión indefinida* $(a_0, a_1, a_2, ..., a_n, ...)$.

DEMOSTRACIONES

Como ya sabemos que $\mathbb{N}$ y $\mathbb{R}$ son infinitos, sólo hemos de probar que $\mathbb{Z}$ y $\mathbb{Q}$ son numerables, que $\mathbb{R}$ no es numerable y que $\mathbb{C}$ tiene la potencia del continuo:

1. El conjunto $\mathbb{Z}$, de los números enteros. es infinito numerable, ya que la siguiente aplicación f es una biyección:

$$f:\mathbb{N}\to\mathbb{Z}, \quad n\mapsto f(n) = \begin{cases} -n/2, & \text{si } n \text{ es par} \\ (n+1)/2, & \text{si } n \text{ es impar} \end{cases}$$

2. El conjunto $\mathbb{Q}$, de los números racionales, es infinito numerable, ya que sus elementos pueden ponerse formando una sucesión indefinida, como pasamos a comprobar. Para cualquier racional $p/q \neq 0$ (en forma irreducible), sea $s > 1$ el número natural $s = |p| + |q|$; cada s proviene de varios p/q, que los escribiremos en el siguiente orden:

$$\frac{s-1}{1}, \frac{s-2}{2}, ..., \frac{2}{s-2}, \frac{1}{s-1}, \frac{-1}{s-1}, \frac{-2}{s-1}, ..., \frac{-(s-2)}{2}, \frac{-(s-1)}{1}$$

(de aquí se quitan aquellos que no estén en forma irreducible).

Pongamos en primer lugar el racional 0, después ponemos los racionales que tienen $s = 2$ (esto es: $1/1, -1/1$), después escribimos los racionales que tienen $s = 3$ (esto es: $2/1, 1/2, -1/2, -2/1$), después aquellos que tienen $s = 4$ (esto es: $3/1, 1/3, -1/3, -3/1$); siguiendo de este modo indefinidamente, se escriben todos los racionales sin excepción en forma de sucesión infinita, como se deseaba.

3. El conjunto $\mathbb{R}$, de los números reales, no es numerable. Para comprobar este aserto recurriremos a la representación decimal ilimitada de los

números reales. Como $\mathbb{R}$ es equipotente a su intervalo $]0, 1[$, habrá que comprobar que cualquier aplicación $f: \mathbb{N} \to]0, 1[$ es no sobreyectiva. En efecto, para cualquiera que sea la aplicación $f: \mathbb{N} \to]0, 1[$, recurramos a la representación decimal ilimitada de $f(n)$, para $n \in \mathbb{N}$,

$$f(n) = 0, a_{n1}a_{n2}...a_{np}...$$

(a_{np} = cifra decimal de lugar p del número real $f(n)$) y, a partir del conocimiento de todas las imágenes $f(n)$, echemos mano del número real

$$r = 0, a_1a_2...a_p... \quad \text{donde} \quad a_1 \neq a_{11},\ a_2 \neq a_{22},\ ...,\ a_p \neq a_{pp},\ ...$$

(por ejemplo, $a_p = 5$ *si* $a_{pp} \neq 5$ y $a_p = 4$ si $a_{pp} = 5$). Es evidente que $r \neq f(n)$ para todo $n \in \mathbb{N}$, ya que r y $f(n)$ difieren, al menos en sus cifras de lugar n, que valen a_n y a_{nn}, respectivamente. Como $r \in]0, 1[$, nos encontramos con que r es un elemento de $]0, 1[$ que no es imagen, por f, de ningún $n \in \mathbb{N}$; esto confirma que f no puede ser sobreyectiva, como se había anunciado.

4. El conjunto $\mathbb{C}$, de los números complejos, tiene la potencia del continuo. Esta afirmación equivale a esta otra: el intervalo real $]0, 1[$ es equipotente al producto cartesiano $]0, 1[\times]0, 1[$, ya que $\mathbb{R}$ es equipotente a $]0, 1[$ y $\mathbb{C}$ es equipotente a $\mathbb{R} \times \mathbb{R}$ y este conjunto lo es, pues, a $]0,1[\times]0, 1[$. Todo se reduce, entonces, a encontrar una aplicación

$$f:]0, 1[\times]0, 1[\to]0, 1[$$

que sea biyectiva.
Recurriendo a la representación decimal ilimitada de los números reales, una tal aplicación biyectiva puede ser la siguiente: para cada $x, y \in]0, 1[$, se toma

$$f(x, y) = 0, a_1b_1a_2b_2...a_nb_n..., \quad \text{donde} \quad \begin{cases} x = 0, a_1a_2...a_n... \\ y = 0, b_1b_2...b_n... \end{cases}$$

RELACION DE ORDEN ENTRE CARDINALES

[380]

Dados dos conjuntos A y B, se pondrá Card $A \leq$ Card B si A es equipotente con una parte de B. De este modo se define una relación binaria entre cardinales que es relación de orden.
Los conjuntos finitos tienen menor cardinal que el de $\mathbb{N}$; esto es, para todo número natural n, es $n <$ Card $\mathbb{N} = \aleph_0$. La potencia del continuo es superior a la potencia de lo numerable, es decir, Card $\mathbb{R} >$ Card $\mathbb{N}$ o sea $\aleph_1 > \aleph_0$.

OBSERVACIONES

1. Dados dos conjuntos A y B, la relación Card $A \leq$ Card B significa que existe una aplicación inyectiva de A en B. Si es Card $A \leq$ Card B y también

Card $A \neq$ Card B, se pondrá Card $A <$ Card B; esto último ocurrirá, pues, si hay alguna aplicación inyectiva de A en B y ninguna que sea biyectiva.

2. Dados dos cardinales $\mathfrak{a} =$ Card A y $\mathfrak{b} =$ Card B, se dice que es $\mathfrak{a} \leqslant \mathfrak{b}$ si Card $A \leqslant$ Card B. De este modo se define una relación binaria entre cardinales, pues ella no depende de los conjuntos A y B; esto es así ya que si $A \simeq A'$ y $B \simeq B'$, entonces Card $A \leqslant$ Card B equivale a Card $A' \leqslant$ Card B'. Esta relación $\leqslant$ es una ordenación entre cardinales, ya que es reflexiva (se comprueba trivialmente), es simétrica (como prueba el *teorema de Berstein*; véase [425], NOTA 9, *c*) y es transitiva (se comprueba trivialmente).
3. No existe ningún conjunto infinito con potencia inferior a la de $\mathbb{N}$, ya que todo conjunto infinito incluye a un conjunto infinito numerable (ver la nota 9, b). Si A es un conjunto infinito, se verifica entonces que $\aleph_0 \leqslant$ Card A.
4. Según se comprobó en el apartado anterior, ninguna de las aplicaciones inyectivas de $\mathbb{N}$ en $\mathbb{R}$ es sobreyectiva; por ello es Card $\mathbb{N} <$ Card $\mathbb{R}$, o sea, $\aleph_0 < \aleph_1$.
5. No se conoce ningún conjunto A tal que Card $\mathbb{N} <$ Card $A <$ Card $\mathbb{R}$, es decir, no se conoce ningún cardinal $\mathfrak{a}$ tal que $\aleph_0 < \mathfrak{a} < \aleph_1$. Se ha probado que la proposición «existe un conjunto A tal que Card $\mathbb{N} <$ Card $A <$ Card $\mathbb{R}$» es indecidible (véase [419], NOTA 3, *b*). Se llama *hipótesis del continuo* al axioma que establece que no existe ningún cardinal estrictamente comprendido entre $\mathbb{N}$ y $\mathbb{R}$.
6. Dado un cardinal cualquiera, siempre existe otro cardinal mayor que él, como prueba el *teorema de Cantor* (véase [425], NOTA 9, *d*), que asegura que, para cualquier conjunto A, es Card $A <$ Card $\mathcal{P}(A)$.

APENDICE

7

Estructuras algebraicas

El álgebra, que antaño sólo se dedicaba a hacer cuentas con números, también se dedica hoy día a operar con otros muchos elementos matemáticos. Dicho de otro modo, el álgebra ha ampliado notablemente su campo de acción por doble motivo: en primer lugar, de manejar sólo números ha pasado a ocuparse de todo tipo de objetos matemáticos y, por otra parte, lo que antes sólo era sumar y multiplicar, se ha convertido ahora en operaciones o leyes de composición cualesquiera. Así, pues, de los campos numéricos, se ha pasado a las estructuras matemáticas abstractas, las cuales se distinguen unas de otras por los axiomas o propiedades que se exigen a las operaciones que en ellas se consideran, perdiendo, así, su importancia tanto la naturaleza concreta de los objetos con los que se opera, como el modo particular de definir las operaciones.

OPERACIONES INTERNAS

[381]

Dado un conjunto $A \neq \emptyset$, se llama *operación interna* o ley de composición interna definida en A a toda aplicación $*$, del siguiente tipo:

$$A \times A \xrightarrow{*} A, \quad (x, y) \longmapsto x * y$$

Se dice que $x * y$ es el resultado de operar x con y.

Un conjunto A y unas operaciones $*$, ..., $\circ$ definidas en A forman lo que se llama una *estructura algebraica*, que se denota poniendo $(A, *, ..., \circ)$. El tipo o clase de estructura se caracterizará atendiendo a las propiedades que satisfagan los elementos de A y a las operaciones definidas en A.

OBSERVACIONES

1. Según la anterior definición, si $*$ es una operación interna en un conjunto A, entonces para cualesquiera que sean $x, y \in A$ existe el elemento $x * y \in A$ y éste es único.
2. Además del signo $*$, para denominar a las operaciones también se suelen usar otros signos como $+$, $\cdot$, $\perp$, $\top$, $\circ$, $\times$, $\wedge$, $\vee$, $\triangle$, $\triangledown$, etc. Si se utiliza el signo «más» $(+)$ o el signo «por» $(\cdot)$, se dice que la notación es aditiva o multiplicativa, respectivamente.

3. En una estructura algebraica $(A; *, ..., \circ)$. en la que A se llama conjunto soporte, al analizarla desde el punto de vista abstracto, interesan las propiedades que la caracterizan, pero no importan ni la naturaleza concreta de los elementos que forman A ni el modo particular de definir las operaciones.
4. Si $(A, *)$ es una estructura algebraica, se dirá que $C \subset A$ es una *parte estable* para la operación $*$ si, para cualesquiera $x, y \in C$, se verifica que $x * y$ también es de C. En tal caso, $(C, *)$ es una estructura algebraica; se dice que $(C, *)$ es una *subestructura* de $(A, *)$ si, además, la estructura $(C, *)$ goza de las propiedades que caracterizan a la estructura $(A, *)$.
5. Sean dados tres conjuntos A, B y C. A una aplicación

$$A \times B \to C, \quad (x, y) \mapsto x * y$$

se la llama *operación externa* o ley de composición externa. Tienen especial interés las operaciones del tipo $A \times B \to B$ (por ejemplo: el producto de escalar por vector, que es un vector) y las del tipo $A \times A \to B$ (por ejemplo: el producto escalar de dos vectores, que es un escalar).

EJEMPLOS

1. La suma y el producto usuales de los campos numéricos $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son operaciones internas para cada uno de ellos; de hecho, estas operaciones son el modelo en el que se apoyó la generalización que ha conducido a la noción de ley de composición. El conjunto $\mathbb{R}_+ = \{x \in \mathbb{R} / x \geq 0\}$ es una parte de $\mathbb{R}$ que es estable para el producto; sin embargo, el conjunto $\mathbb{R}_- = \{x \in \mathbb{R} / x \leq 0\}$ no es parte estable para el producto.
2. Para el conjunto de las partes de un conjunto, la unión y la intersección son operaciones internas.
3. Para el conjunto $A = \mathcal{F}(X, X)$, de las aplicaciones de X en X, la composición de aplicaciones es una operación interna. Los conjuntos de las aplicaciones inyectivas, sobreyectivas y biyectivas de X en X son tres partes estables de A respecto de la composición de aplicaciones.

EJERCICIO

Si x e y son dos números reales y siempre que la siguiente expresión tenga sentido, llamemos $x * y$ a:

$$x * y = \frac{x + y}{x^2 + y^2}$$

Analícese si $*$ es una operación interna para cada uno de los siguientes conjuntos:

$$\mathbb{R}, \ R^* = \mathbb{R} - \{0\}, \quad \mathbb{R}_+^* = \{x \in \mathbb{R} / x > 0\} \quad \text{y} \quad \mathbb{R}_-^* = \{x \in \mathbb{R} / x < 0\}$$

RESOLUCION

Como no existe $0*0$, resulta que $*$ no es operación para $\mathbb{R}$. Para $\mathbb{R}^*$, tampoco $*$ es operación ya que $x*(-x)=0$ no pertenece a $\mathbb{R}^*$. Para $\mathbb{R}_+^*$ y para $\mathbb{R}_-^*$ es claro que $*$ es operación interna.

[382]

Si $*$ es una operación interna, definida en un conjunto A, se dice que:

- $*$ es *asociativa* si $(x*y)*z=x*(y*z)$, $\forall x, y, z \in A$.
- $*$ es *conmutativa* si $x*y=y*x$, $\forall x, y \in A$.
- $*$ es *distributiva* respecto de otra operación interna $\circ$ de A si $x*(y\circ z)=(x*y)\circ(x*z)$ e $(y\circ z)*x=(y*x)\circ(z*x)$, $\forall x, y, z \in A$.
- $e \in A$ es *elemento neutro* si $x*e=e*x=x$, $\forall x \in A$.

  Si en $(A; *)$ hay un elemento neutro, entonces éste es único:
- $x \in A$ es *simetrizable* si existe $x' \in A$ tal que $x*x'=x'*x=e$ (aquí se supone que existe el neutro e); entonces, x' se llama simétrico de x.

  Si $*$ es asociativa y x es simetrizable, entonces x tiene un sólo simétrico
- $x \in A$ es *regular* o simplificable si, $\forall y_1, y_2 \in A$, se verifica que

$$x*y_1=x*y_2 \Rightarrow y_1=y_2 \quad \text{e} \quad y_1*x=y_2*x \Rightarrow y_1=y_2$$

Si $*$ es asociativa, todo elemento simetrizable de A es regular.

DEMOSTRACIONES

1. Comprobemos primeramente la unicidad del elemento neutro: si e y f fuesen, ambos, elementos neutros, resultaría que $e*f=e$ (pues f es neutro) y $e*f=f$ (pues e es neutro), luego habría de ser $e=f$.
2. Veamos ahora que, si $*$ es asociativa, un elemento $x \in A$ no puede tener más de un simétrico. En efecto: si $x', x'' \in A$ fuesen simétricos de $x \in A$, habría de ser $x'=x''$ ya que:

$$x'*x*x''=\begin{cases}(x'*x)*x''=e*x''=x''\\ x'*(x*x'')=x'*e=x'\end{cases}$$

3. Comprobemos, finalmente, que si $*$ es asociativa todo elemento $x \in A$ simetrizable es regular. Así ocurre ya que, llamando $x' \in A$ al simétrico de A, para $y_1, y_2 \in A$ es:

$$x*y_1=x*y_2 \Rightarrow x'*x*y_1=x'*x*y_2 \Rightarrow e*y_1=e*y_2 \Rightarrow y_1=y_2$$
$$y_1*x=y_2*x \Rightarrow y_1*x*x'=y_2*x*x' \Rightarrow y_1*e=y_2*e \Rightarrow y_1=y_2$$

OBSERVACIONES

1. Si $*$ es una operación asociativa, en lugar de poner $(x*y)*z$ o $x*(y*z)$, que son iguales, se simplifica la notación escribiendo $x*y*z$. Si $*$ es

asociativa en A y dados $x_1, x_2, ..., x_n \in A$, cualquiera que sea la forma de agruparlos, para operar con ellos de dos en dos en el orden dado, se llega a un mismo resultado, que se denota poniendo $x_1 * x_2 * \cdots * x_n$; es decir, el resultado de la operación no depende de cómo se coloquen los paréntesis.

2. El elemento $x * x * \overset{n}{\cdots} * x$ (resultado de operar n veces con x) se denota: *a*) si la notación es aditiva, poniendo nx (múltiplo de x); b) si la notación es multiplicativa x^n (potencia de x).
3. Cuando la notación es aditiva, al elemento neutro se le llama *elemento nulo* ($e = o$) y al simétrico de un elemento x se le llama el *opuesto* de x, poniendo $x' = -x$. En notación multiplicativa, al neutro se le llama *elemento unidad* ($e = 1$) y al simétrico de un elemento x se le llama el *inverso* de x, poniendo $x' = x^{-1}$.
4. Si $(A, *)$ es una estructura algebraica con elemento neutro y si $*$ es asociativa, dados $a, b \in A$ donde a tiene simétrico a', las ecuaciones $x * a = b$ y $a * x = b$ ($x \in A$ es la incógnita) tienen solución única, que es $x = b * a'$ y $x = a' * b$, respectivamente.
5. En una estructura algebraica $(A, *)$ se dice que $x \in A$ es un elemento *idempotente* si $x * x = x$.

EJEMPLOS

1. La suma y el producto usuales de los campos numéricos $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son operaciones asociativas y conmutativas; el producto es distributivo respecto de la suma, pero la suma no lo es respecto del producto.
2. Para el conjunto $A = \mathcal{P}(U)$, de las partes de un conjunto U dado, la unión y la intersección son asociativas y conmutativas; cada una de ellas es distributiva respecto de la otra. Para la intersección hay elemento neutro que es U y, salvo U, ningún otro elemento admite simétrico.
3. Para el conjunto $A = \mathcal{F}(X, X)$, de las aplicaciones de X en X, con la operación $\circ$ (composición de aplicaciones), la aplicación idéntica es el elemento unidad. Una aplicación $f \in A$ tiene simétrica si y sólo si es biyectiva y, entonces, su simétrica es su aplicación recíproca f^{-1}.
4. En $\mathbb{N}$, considérese la operación interna $*$ definida por:

$$n * m = n(n + m)$$

Esta operación no es asociativa ya que, por ejemplo, $(1 * 2) * 1 = 3 * 1 = 12$ es distinto de $1 * (2 * 1) = 1 * 6 = 7$. Esta operación tampoco es conmutativa ya que, por ejemplo $1 * 2 = 3$ es distinto de $2 * 1 = 6$.

EJERCICIOS

1. Sea $(A, *)$ una estructura algebraica con elemento unidad e y en la que $*$ es asociativa. Si $x, y \in A$ tienen elementos simétricos $x', y' \in A$, compruébese que $x * y$ también tiene simétrico y que éste es $(x * y)' = y' * x'$.
2. Considérese la estructura algebraica $(\mathbb{R}, *)$, siendo

$$x * y = x + (y - 1)|x|$$

Compruébese que $1 \in \mathbb{R}$ es el elemento neutro, que todo $x > 0$ admite por simétrico a $x' = 1/x$ y que los $x \leq 0$ carecen de simétrico.

RESOLUCION

1. Todo se reduce a comprobar que operando $x * y$ con $y' * x'$ se obtiene el elemento neutro. Así es, ya que:

$$(x * y) * (y' * x') = x * (y * y') * x' = x * e * x' = x * x' = e$$
$$(y' * x') * (x * y) = y' * (x' * x) * y = y' * e * y = y' * y = e$$

2. El elemento 1 es el neutro ya que:

$$1 * x = 1 + (x - 1) \cdot 1 = x \quad \text{y} \quad x * 1 = x + (1 - 1)|x| = x$$

Para $x \geq 0$ el simétrico de x es $1/x$ ya que:

$$x * \frac{1}{x} = x + \left(\frac{1}{x} - 1\right)x = 1 \quad \text{y} \quad \frac{1}{x} * x = \frac{1}{x} + (x - 1)\frac{1}{x} = 1$$

Para $x < 0$ no existe simétrico de x, ya que:

$$x * x' = 1 \Rightarrow x(2 - x') = 1 \Rightarrow x' = 2 - \frac{1}{x} > 0 \Rightarrow x' * x = 2x - 1$$

y este último elemento no puede valer 1 (es $x < 0$). Para $x = 0$, es $x * y = 0 \neq 1$ para todo $y \in \mathbb{R}$, luego no hay simétrico de x.

OPERACION COMPATIBLE CON UNA EQUIVALENCIA

Una operación interna y una relación de equivalencia, ambas definidas entre elementos de un mismo conjunto, se dicen compatibles si el resultado de operar con dos elementos dados, cualesquiera, es equivalente al resultado de operar con elementos equivalentes a los dados; más exactamente.

Esto es lo que ocurre con el siguiente caso: En el conjunto $\mathbb{C}$ de los números complejos considérense la operación producto y la equivalencia $\sim$ según la que $z \sim z'$ significa que $|z| = |z'|$; éstas, operación y equivalencia, son compatibles, ya que:

$$\left.\begin{array}{lll} z \sim z' & \Rightarrow & |z| = |z'| \\ w \sim w' & \Rightarrow & |w| = |w'| \end{array}\right\} \Rightarrow |zw| = |z'w'| \Rightarrow [zw \sim z'w']$$

[383]

Sea A un conjunto en el que hay definidas una relación de equivalencia $\sim$ y una operación interna $*$. Se dice que $\sim$ y $*$ son compatibles si, para cualesquiera $x, x', y, y' \in A$, se verifica:

$$(x \sim x', y \sim y') \Rightarrow x * y \sim x' * y'$$

Teorema. Si las anteriores $\sim$ y $*$ son compatibles, entonces

$$(A/\sim) \times (A/\sim) \to A/\sim \qquad ([x], [y]) \mapsto [x * y]$$

es una aplicación; es decir, de este modo se define una operación interna en el conjunto cociente $A/\sim$, que se denotará con el mismo signo $*$. A ella se le llama *operación cociente* y queda, pues, definida por:

$$[x] * [y] = [x * y]$$

DEMOSTRACION

Como todo par de clases $([x], [y])$ tiene imagen, pues existe la clase $[x * y]$, sólo hay que comprobar que esta imagen es única, es decir, que ella no depende de los representantes x e y que se tomen para definir las clases $[x]$ e $[y]$, o sea, que, si $[x] = [x']$ e $[y] = [y']$, entonces $[x * y] = [x' * y']$. Así es, en efecto, pues como $\sim$ y $*$ son compatibles, se tiene:

$$\left.\begin{array}{l} [x] = [x'] \Leftrightarrow x \sim x' \\ [y] = [y'] \Leftrightarrow y \sim y' \end{array}\right\} \Rightarrow x * y \sim x' * y' \Rightarrow [x * y] = [x' * y']$$

OBSERVACION

Dadas una estructura algebraica $(A, *)$ y una equivalencia $\sim$ entre elementos de A que sea compatible con $*$, en general se verifica que la estructura cociente $(A/\sim, *)$ «hereda» las propiedades de la estructura $(A, *)$. Más exactamente: si $*$ es asociativa o conmutativa en A, entonces también lo es en $A/\sim$; si e es el neutro de $(A, *)$, entonces $[e]$ es el neutro de $(A/\sim, *)$; si $x \in A$ tiene por simétrico a $x' \in A$ en $(A, *)$, entonces $[x']$ es la clase simétrica de la $[x]$ en $(A/\sim, *)$.

EJEMPLO

Sea $p > 1$ un número natural dado. En $\mathbb{Z}$, la relación binaria $\sim$ definida por

$$n \sim n' \Leftrightarrow (n - n' \text{ es múltiplo de } p)$$

es una equivalencia, que se llama *congruencia* módulo p y se acostumbre a denotar poniendo $n \equiv n'(p)$ (véase el ejemplo 2 de [368]). Las clases de esta equivalencia están formadas por todos los enteros que dan un mismo resto al dividirlos por p, esto es, los conjuntos:

$$[r] = \{r + kp / k \in \mathbb{Z}\} \quad \text{para } r = 0, 1, 2, \ldots, p-1$$

Nótese que

$$\left.\begin{array}{l} n \sim n' \Rightarrow n' = n + kp \\ m \sim m' \Rightarrow m' = m + hp \end{array}\right\} \Rightarrow n' + m' = (n+m) + (k+h)p \Rightarrow n + m \sim n' + m'$$

(donde k y h son ciertos números enteros); es decir, la equivalencia $\sim$ es compatible con la suma $(+)$ en $\mathbb{Z}$. La operación cociente en $\mathbb{Z}/\sim$, que se denota poniendo $\mathbb{Z}/\sim = \mathbb{Z}/(p)$, está dada por:

$$[n] + [m] = [n + m]$$

HOMOMORFISMOS

Para señalar que dos estructuras algebraicas son esencialmente análogas, que sus diferencias lo son de apariencia pero no de fondo, se dice que son homomorfas (semejantes en la forma). Esta idea se concreta recurriendo a una aplicación entre las estructuras que «respete» la operación, es decir, tal que al resultado de operar con dos elementos de una de las estructuras le corresponda, en la otra, el resultado de operar con las imágenes de los elementos de partida.

[384]

Dadas dos estructuras algebraicas $(A, *)$ y $(B, \divideontimes)$, se dice que una aplicación $f: A \to B$ es lineal para las operaciones $*$ y $\divideontimes$ o que es un *homomorfismo* de $(A, *)$ en $(B, \divideontimes)$, y se pondrá $f: (A, *) \to (B, \divideontimes)$, si para cualesquiera $x, x' \in A$ se verifica:

$$f(x * x') = f(x) \divideontimes f(x)$$

Composición de homomorfismo: supóngase que $f: (A, *) \to (B, \divideontimes)$ y $g: (B, \divideontimes) \to (C, \otimes)$ son dos homomorfismos; entonces su composición $g \circ f: A \to C$ es un homomorfismo de la estructura $(A, *)$ en la $(C, \otimes)$.

DEMOSTRACION

Para cualesquiera que sean $x, x' \in A$ y por ser f un homomorfismo, será:

$$f(x * x') = f(x) \divideontimes f(x')$$

Como $y = f(x)$ e $y' = f'(x)$ pertenecen a B y por ser g un homomorfismo, será

$$g(y \divideontimes y') = g(y) \otimes g(y')$$

Consecuentemente, resulta que:

$$(g \circ f)(x * x') = g(f(x * x')) = g(f(x) \circledast f(x')) = g(f(x)) \otimes g(f(x')) =$$
$$= [(g \circ f)(x)] \otimes [(g \circ f)(x')]$$

por lo cual, $g \circ f$ es un homomorfismo de $(A;\ *)$ en $(C;\ \otimes)$.

EJEMPLOS

1. Sea A el conjunto de las funciones reales que tienen integral en el intervalo $[0, 1]$. La suma usual de funciones es una operación interna para A (o sea, la suma de funciones con integral es otra función que también es integrable) y se verifica además (según se sabe de cálculo integral) que:

$$\int_0^1 (\varphi + \psi) = \int_0^1 \varphi + \int_0^1 \psi$$

Por tanto la aplicación

$$F: A \to \mathbb{R}, \quad \varphi \mapsto F(\varphi) = \int_0^1 \varphi$$

es un homomorfismo de $(A, +)$ *en* $(\mathbb{R}, +)$. Sin embargo, F no es un homomorfismo de $(A, \cdot)$ en $(\mathbb{R}, \cdot)$ ya que, en general,

$$\int_0^1 (\varphi \cdot \psi) \neq \left(\int_0^1 \varphi\right) \cdot \left(\int_0^1 \psi\right)$$

2. La aplicación $f: \mathbb{R}_+^* \to \mathbb{R}$ que a cada número real $x > 0$ le hace corresponder $f(x) = Lx$ (logaritmo neperiano) es un homomorfismo de la estructura $(\mathbb{R}_+^*, \cdot)$ en la $(\mathbb{R}, +)$, ya que

$$f(x \cdot x') = L(x \cdot x') = Lx + Lx' = f(x) + f(x')$$

La aplicación $g: \mathbb{R} \to \mathbb{C}$ que a cada real y le hace corresponder el número complejo

$$g(y) = \cos y + i \operatorname{sen} y = e^{iy}$$

es un homomorfismo de $(\mathbb{R};\ +)$ en $(\mathbb{C};\ \cdot)$, ya que

$$g(y + y') = e^{i(y+y')} = e^{iy} \cdot e^{iy'} = g(y) \cdot g(y')$$

La composición de f con g, esto es,

$$g \circ f: \mathbb{R}_+^* \to \mathbb{C}, \quad x \mapsto g(f(x)) = \cos(Lx) + i \operatorname{sen}(Lx)$$

es, pues, un homomorfismo de $(\mathbb{R}_+^*, \cdot)$ en $(\mathbb{C}, \cdot)$.

ISOMORFISMOS

[385]

Si $f:(A, *)\to(B, \bar{*})$ es un homomorfismo, entonces $f(A)$ es una parte estable de B respecto de $\bar{*}$, es decir, $(f(A), \bar{*})$ es una estructura algebraica; a ella se la llama *imagen homomorfa* de $(A, *)$ por f.

Un homomorfismo $f:(A, *)\to(B, \bar{*})$ se dice que es un *isomorfismo* si $f:A\to B$ es biyectiva y, en tal caso, f^{-1} también es un isomorfismo; se dice entonces que $(A, *)$ y $(B, \bar{*})$ son estructuras isomorfas. Dos estructuras isomorfas son identificables, se puede considerar que son dos concreciones particulares de una misma estructura abstracta. La composición de dos isomorfismos es otro isomorfismo.

Un homomorfismo de una estructura en sí misma, $f:(A, *)\to(A, *)$, recibe el nombre de *endomorfismo*. Los endomorfismos biyectivos se llaman *automorfismos*.

DEMOSTRACION

Para comprobar que $f(A)$ es una parte estable de B respecto de la operación $\bar{*}$, hay que cerciorarse de que, para cualesquiera $y, y'\in f(A)$, el elemento $y\bar{*}y'$ también es de $f(A)$. Así es ya que: como se sabe que $y=f(x)$ e $y'=f(x')$ para algunos $x, x'\in A$, y dado que f es un homomorfismo, se tiene que:

$$y\bar{*}y'=f(x)\bar{*}f(x')=f(x*x')$$

y este último elemento es de $f(A)$ puesto que $x*x'\in A$.

OBSERVACIONES

1. Si $f:(A, *)\to(B, \bar{*})$ es un homomorfismo, como $(f(A), \bar{*})$ es una estructura algebraica, resulta que $f:(A, *)\to(f(A), \bar{*})$ es un homomorfismo inyectivo. Respecto del homomorfismo dado, el conjunto $B-f(A)$ desempeña un papel secundario.
2. Si $f:(A, *)\to(B, \bar{*})$ es un homomorfismo, las propiedades de la estructura $(A; *)$ se «trasladan», en general, a su imagen homomorfa $(f(A), \bar{*})$, pero de ellas no goza necesariamente la estructura de llegada $(B, \bar{*})$. En concreto:

   - Si $*$ es asociativa o conmutativa en A, entonces $\bar{*}$ es asociativa o conmutativa, respectivamente, en $f(A)$, pero puede no serlo en B.
   - Si e es el elemento neutro de $(A, *)$, entonces $f(e)$ es elemento neutro de $(f(A), \bar{*})$, pero puede no serlo de $(B, \bar{*})$.

- Si en $(A, *)$ los elementos x y x' son simétrico uno de otro, entonces $f(x)$ y $f(x')$ son simétrico uno de otro en $(f(A), \divideontimes)$, pero pueden no serlo en $(B, \divideontimes)$.

3. Dadas una estructura algebraica $(A, *)$ y una aplicación biyectiva $f: A \to B$, al conjunto B se le puede dotar de una operación interna $\divideontimes$ tal que $(B, \divideontimes)$ sea una estructura isomorfa a $(A, *)$, de manera que f sea un isomorfismo entre ellas. Esto se logra llamando:

$$y \divideontimes y' = f(f^{-1}(y) * f^{-1}(y')), \quad \text{para } y, y' \in B$$

Se dice que, con ello, se traslada o transporta la estructuración algebraica de A al conjunto B.

EJEMPLO

Considérense las estructuras algebraicas $(\mathbb{R}, \cdot)$ y $(\mathbb{R}^2, *)$, donde $\cdot$ es el producto usual y $*$ es la operación definida por:

$$(x, y) * (x', y') = (x \cdot x', y \cdot y')$$

La aplicación $f: \mathbb{R} \to \mathbb{R}^2$, $f(x) = (x, 0)$ es un homomorfismo de $(\mathbb{R}, \cdot)$ en $(\mathbb{R}^2, *)$ ya que

$$f(x_1 \cdot x_2) = (x_1 \cdot x_2, 0) = (x_1, 0) * (x_2, 0) = f(x_1) * f(x_2)$$

La imagen de la aplicación f es $I = \{(x, 0)/x \in \mathbb{R}\}$ y la imagen homomorfa del homorfismo f es $(I, *)$. El número 1 es la unidad de $(\mathbb{R}, \cdot)$ y $f(1) = (1, 0)$ es el elemento neutro de $(I, *)$; nótese que $(\mathbb{R}^2, *)$ tiene a $(1, 1)$ como elemento neutro y que éste es distinto del neutro $(1, 0)$ de $(I, *)$.

APENDICE

8 Grupos

A los grupos, como a las demás estructuras algebraicas, se llega, mediante un proceso de abstracción, a partir de diversas teorías matemáticas concretas en las que se dispone de una operación interna, entre unos ciertos objetos matemáticos, que goza de unas determinadas propiedades. Estas darán lugar a los axiomas de la teoría de grupos; en ella poco importa la naturaleza de los objetos matemáticos que se consideren o cuál es la operación, con tal de que se satisfagan los axiomas.

Para que podamos apreciar mejor en qué consiste la generalización que conduce al concepto de grupo, empezaremos hablando de algún ejemplo sobre el que apoyarla.

EJEMPLO

Vamos a considerar aquí las aplicaciones biyectivas $f: A \to A$, de un cierto conjunto dado $A \neq \emptyset$ en sí mismo. Según sabemos, la composición de dos cualesquiera de estas biyecciones es, también, una biyección de A en A o, dicho de otro modo, la composición de aplicaciones ($\circ$) es una operación interna para el conjunto F de las biyecciones de A en A. Así, pues, el par $(F, \circ)$ es una estructura algebraica.

De acuerdo con lo que ya sabemos acerca de las biyecciones, resulta que, para la estructura $(F, \circ)$, se verifican las siguientes propiedades:

- La operación $\circ$ es asociativa; esto es, para cualesquiera que sean las aplicaciones $f, g, h \in F$, se cumple que $h \circ (g \circ f) = (h \circ g) \circ f$.
- La aplicación idéntica $i: A \to A$ (recuérdese que, para todo $a \in A$, es $i(a) = a$) es elemento neutro para la composición de aplicaciones; es decir, $f \circ i = i \circ f = f$ para cualquier $f \in F$.
- Toda biyección $f \in F$ tiene biyección recíproca $f^{-1}: A \to A$; o sea, para cada $f \in F$ existe un $f^{-1} \in F$ tal que $f^{-1} \circ f = f \circ f^{-1} = i$.

Pues bien, para expresar que se verifican estas tres propiedades, se dice que la estructura algebraica $(F, \circ)$ es un grupo. Así, pues, un grupo es una estructura algebraica, con una sola operación, caracterizada por los siguientes axiomas: la operación es asociativa, existe elemento neutro y cada elemento tiene su elemento simétrico.

Nótese que, en el anterior ejemplo, la operación del grupo no es conmutativa (en general $f \circ g$ y $g \circ f$ son distintas); cuando la operación sea conmutativa, se dirá que el grupo es conmutativo o abeliano.

CONCEPTO DE GRUPO

[386]

Un conjunto G y una operación interna $*$ definida en él (*) se dice que forman un *grupo* $(G, *)$ si se verifican las siguientes propiedades (axiomas de grupo):

- $(a*b)*c = a*(b*c)$ para cualesquiera $a, b, c \in G$ (asociatividad).
- Existe un $e \in G$ tal que $a*e = e*a = a$, para todo $a \in G$ (e se llama *elemento neutro*).
- Para cada $a \in G$ existe un $a' \in G$ tal que $a*a' = a'*a = e$ (a' se llama *elemento simétrico* de a).

Se dice que $(G, *)$ es un *grupo abeliano* o conmutativo si, además, se verifica la siguiente propiedad:

- $a*b = b*a$, para cualesquiera $a, b \in G$ (conmutatividad).

(*) Así, pues, para cualesquiera que sean $a, b \in G$, existe $a*b$ y éste es también un elemento de G.

OBSERVACIONES

- Cuando no se haga necesario recordar cuál es la operación del grupo, en lugar de representarle por $(G, *)$, se le suele denotar por G, omitiendo el signo $*$.
- Normalmente, la operación del grupo se denota poniendo $\cdot$ (signo «por» del producto); con esta «notación multiplicativa», al elemento netro se le suele llamar unidad (1) y al simétrico de un $a \in G$ se le llama su «inverso» y se denota por a^{-1}. Cuando así se hace, para representar un producto $a \cdot b$ se simplifica la notación y se pone simplemente ab.
- Cuando la operación se representa por un $+$ (signo «más» de la suma), se dice que la notación es «aditiva»; con ella, el elemento neutro se llama «cero» (0) o «nulo» y el elemento simétrico de a se denota poniendo $-a$, que se llama «opuesto» de a.

PRIMERAS PROPIEDADES

[387]

En todo grupo $(G, *)$, se verifican las siguientes propiedades:

1.ª El elemento neutro, e, es único.
2.ª Cada elemento $a \in G$ tiene un sólo simétrico $a' \in G$.
3.ª $(a*b)' = b'*a'$, para cualesquiera $a, b \in G$.
4.ª $(a')' = a$, para cualquier elemento $a \in G$.
5.ª Todos los elementos de G son regulares o simplificables.
6.ª Las ecuaciones $a*x = b$ y $x*a = b$ (donde $a, b \in G$ son datos; $x \in G$ es la incógnita) tienen solución única, que es $x = a'*b$ y $x = b*a'$, respectivamente.

DEMOSTRACION

Salvo las propiedades 4.ª y 6.ª, todas las demás ya se han demostrado, a lo largo de [382], cuando hablábamos de operaciones internas en general. Para las propiedades 4.ª y 6.ª, se tiene:

- El elemento $(a')'$ es, por definición, tal que $a' * (a')' = e$; por tanto, operando, por delante, con a, se tiene, sucesivamente, que:

$$a*(a'*(a')') = a*e, \quad (a*a')*(a')' = a, \quad e*(a')' = a, \quad (a')' = a$$

- Para la última propiedad, operando con a' se obtiene que:

$$\begin{aligned} a*x = b &\Leftrightarrow a'*(a*x) = a'*b \Leftrightarrow (a'*a)*x = a'*b \Leftrightarrow \\ &\Leftrightarrow e*x = a'*b \Leftrightarrow x = a'*b \\ x*a = b &\Leftrightarrow (x*a)*a' = b*a' \Leftrightarrow x*(a*a') = b*a' \Leftrightarrow \\ &\Leftrightarrow x*e = b*a' \Leftrightarrow x = b*a' \end{aligned}$$

EJEMPLOS

1. Los conjuntos numéricos $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ son grupos aditivos, esto es, respecto de la operación suma; todos ellos son grupos abelianos.
2. Los conjuntos $\mathbb{Q}^*$, $\mathbb{R}^*$ y $\mathbb{C}^*$(que son los que resultan de prescindir del 0 en $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$) son grupos multiplicativos, es decir, respecto de la operación producto usual; estos grupos son abelianos. Los conjuntos $\mathbb{Q}_+^*$ y $\mathbb{R}_+^*$ (números estrictamente positivos) son grupos multiplicativos.
3. Sea dado un triángulo equilatero ABC; sea O su centro. Consideremos los siguientes seis movimientos planos: las tres simetrías S_A, S_B y S_C respecto de los ejes OA, OB y OC y los tres giros $G_0G_{120}G_{240}$ de centro O y amplitudes de 0°, 120° y 240°. Estos, que son los seis movimientos que dejan invariante al triángulo dado, forman grupo con relación a la operación $\circ$, esto es, respecto de la composición de movimientos. En este grupo, el elemento neutro es el G_0; cada uno de los S_A, S_B y S_C es el elemento simétrico de sí mismo; cada uno de los G_{120} y G_{240} es el simétrico del otro.
4. Si $(G_1, \cdot)$ y $(G_2, \cdot)$ son dos grupos (en ambos hemos utilizado notación multiplicativa), entonces el producto cartesiano $G_1 \times G_2$ es un grupo respecto de la operación $*$ definida mediante

$$(a_1, a_2)*(b_1, b_2) = (a_1 \cdot b_1, a_2 \cdot b_2) \quad \forall a_1, b_1 \in G_1 \quad \text{y} \quad \forall a_2, b_2 \in G_2$$

Si $e_1 \in G_1$ y $e_2 \in G_2$ son los elementos neutros de G_1 y G_2, entonces (e_1, e_2) es el neutro de $G_1 \times G_2$. Si a_1' es el simétrico de $a_1 \in G_1$ y si a_2' es el simétrico de $a_2 \in G_2$, entonces (a_1', a_2') es el simétrico de (a_1, a_2) en $G_1 \times G_2$.

SUBGRUPOS

Dado un grupo $(G, *)$ y si H es un subconjunto cualquiera de G, en general H no será un grupo respecto de la operación $*$, es más, ni siquiera ocurrirá que

$*$ sea operación interna para H. Pues bien, aquellos subconjuntos H de G que resulten ser grupos, con la misma operación $*$ de G, se llaman subgrupos de G. Más exactamente:

[388]

Sea $(G, *)$ un grupo y sea H un subconjunto de G. Se dice que H es un subgrupo del grupo $(G, *)$ si H es una parte estable para la operación $*$ y la estructura $(H, *)$ es un grupo. Esto es:

Definición de subgrupo. Se dice que $H \subset G$ es un subgrupo de $(G, *)$ si:

1. $(a, b \in H) \Rightarrow a * b \in H$ (o sea, $*$ es operación para H).
2. El elemento neutro e de $(G, *)$ pertenece a H.
3. $a \in H \Rightarrow a' \in H$ (por a' se denota al simétrico de a en G).

Caracterización de los subgrupos. El subgrupo H de G es un subgrupo de $(G, *)$ si y sólo si se verifican las dos condiciones siguientes:

i) $H \neq \emptyset$.

ii) $(a, b \in H) \Rightarrow a * b' \in H$.

DEMOSTRACION

- Si H es un subgrupo, será: i) $H \neq \emptyset$ ya que $e \in H$; y ii) si $a, b \in H$, se sabe (por 3) que $b' \in H$ y entonces (por 1), como $a, b' \in H$, resulta que $a * b' \in H$.
- Si H satisface a las condiciones i) e ii), entonces se verifican los tres requisitos de grupo:

  — Como $H \neq \emptyset$, existe $a \in H$, luego (según ii) $a * a' \in H$, esto es, $e \in H$.
  — Si $a \in H$, como $e \in H$, resulta (según ii) que $e * a' \in H$, esto es, $a' \in H$.
  — Si $a, b \in H$, como entonces $b' \in H$, resulta (según ii) que $a * (b')' = a * b \in H$.

EJEMPLOS

1. Un grupo G, cualquiera, tiene como subgrupos (entre otros posibles) al propio G y a $\{e\}$ (e es el elemento neutro). A éstos se les llama «subgrupos triviales»; a los demás se les llama «subgrupos propios».
2. El grupo aditivo $(\mathbb{Z}, +)$ admite como subgrupo al conjunto de los múltiplos de p (entero dado). El grupo multiplicativo $(\mathbb{R}^*, \cdot)$, de los reales no nulos, admite como subgrupo al conjunto $\mathbb{R}_+^*$ de los reales estrictamente positivos. El grupo multiplicativo $(\mathbb{C}^*, \cdot)$, de los complejos no nulos, admite como subgrupo al conjunto de los complejos que tienen módulo unidad.
3. Sea $(G, \circ)$ el grupo de las biyecciones de un conjunto A en sí mismo, donde $\circ$ es la composición de aplicaciones; sea $a_0 \in A$ un elemento fijo. Las biyecciones de A en A que dejan invariante a a_0 forman un subgrupo de G.

EJERCICIO

Dado un grupo $(G, *)$, sea $\{H_i / i \in I\}$ una familia de subgrupos de G. Demuéstrese que la intersección H de los conjuntos H_i es también un subgrupo de G.

RESOLUCION

- En primer lugar, como el elemento neutro $e \in G$ pertenece a todos los H_i, resulta que $e \in H$, luego es $H \neq \emptyset$.
- En el supuesto de que $a, b \in H$, hay que comprobar que $a * b' \in H$. Sabemos que $a, b \in H_i$ para todo $i \in I$ y, como H_i es un subgrupo, resulta que $a * b' \in H_i$ para todo $i \in I$, luego $a * b'$ pertenece a la intersección H de los H_i.

HOMOMORFISMOS DE GRUPOS

El concepto de homomorfismo de grupos tiene su origen en algo tan simple como «la regla de tres». En ésta hay una correspondencia entre dos conjuntos numéricos (en los que se puede sumar) y las cosas funcionan de manera que a la suma de dos números, del conjunto de partida, le corresponde la suma de las cantidades que le corresponden a cada uno de los dos sumandos; esto es, la imagen de la suma es igual a la suma de las imágenes. Si la aplicación en cuestión lo es entre dos grupos, sin más que sustituir las sumas anteriores por las operaciones de los grupos, obtendremos la definición de homomorfismo:

[389]

Dados dos grupos $(G, *)$ y $(H, \circ)$, se dice que una aplicación $f: G \to H$ es un *homomorfismo* del grupo $(G, *)$ en el grupo $(H, \circ)$ si, para cualesquiera que sean $a, b \in G$, se verifica que

$$f(a * b) = f(a) \circ f(b)$$

El homomorfismo $f: (G, *) \to (H, \circ)$ se dice que es un *isomorfismo* si la aplicación $f: G \to H$ es una biyección de G en H (*). Un homomorfismo $\varphi: (G, *) \to (G, *)$, de un grupo en sí mismo, recibe el nombre de *endomorfismo*. Un endomorfismo $\varphi: (G, *) \to (G, *)$ se dice que es un *automorfismo* si la aplicación $\varphi: G \to G$ es biyectiva.

(*) Si el homomorfismo f es inyectivo, se dice que es un monomorfismo; si f es sobreyectivo, se dice que es un epimorfismo.

EJEMPLOS

1. Considérense los grupos $(\mathbb{R}, +)$ y $(\mathbb{R}^*_+, \cdot)$ y la aplicación exponencial:

$$f: \mathbb{R} \to \mathbb{R}^*_+ \quad f(x) = e^x$$

Esta aplicación es un homomorfismo entre los citados grupos, ya que para cualesquiera $x, y \in \mathbb{R}$ es:

$$e^{x+y} = e^x \cdot e^y, \quad \text{es decir} \quad f(x+y) = f(x) \cdot f(y)$$

Como f es una aplicación biyectiva, se trata de un isomorfismo.

2. Sea $\mathcal{F}$ el conjunto de las funciones reales definidas en $\mathbb{R}$; sea $\mathcal{D}$ el conjunto de las funciones reales definidas en $\mathbb{R}$ que son derivables; tanto $\mathcal{F}$ como $\mathcal{D}$ son grupos respecto de la suma usual de funciones; sea $D: \mathcal{D} \to \mathcal{F}$ la aplicación derivada, es decir, aquella que a cada función derivable f le asocia su derivada f', o sea, $D(f) = f'$. Esta aplicación es un homomorfismo del grupo $(\mathcal{D}, +)$ en el grupo $(\mathcal{F}, +)$, ya que

$$(f+g)' = f' + g', \quad \text{es decir, } D(f+g) = D(f) + D(g)$$

PROPIEDADES DE LOS HOMOMORFISMOS

[390]

Si $f: G \to H$ es un homomorfismo del grupo $(G, *)$ en el grupo $(H, \circ)$, se verifica que:

1. $f(e) = \varepsilon$ (donde e y ε son los elementos neutros de G y H).
2. $f(a') = f(a)'$, $\forall a \in G$ (a' y $f(a)'$ son los simétricos de a y $f(a)$).
3. ($\bar{G}$ es un subgrupo de G) $\Rightarrow$ ($f(\bar{G})$ es un subgrupo de H).
4. ($\bar{H}$ es un subgrupo de H) $\Rightarrow$ ($f^{-1}(\bar{H})$ es un subgrupo de G).
5. (f es un isomorfismo) $\Rightarrow$ (f^{-1} es un isomorfismo).

---

6. Si $f: (G, *) \to (H, \triangle)$ y $g: (H, \triangle) \to (J, \star)$ son homomorfismos de grupos, entonces también lo es $g \circ f: (G, *) \to (J, \star)$.

DEMOSTRACION

1. Recurriendo a un elemento cualquiera $a \in G$, obtenemos que $f(a * e) = f(a) * f(e)$, es decir, $f(a) = f(a) * f(e)$. Operando en ambos miembros de esta igualdad con el simétrico de $f(a)$, se obtiene que $\varepsilon = \varepsilon \circ f(e)$, esto es, que $\varepsilon = f(e)$.
2. Para cualquiera que sea $a \in G$, se verifica que

$$f(a) \circ f(a') = f(a * a') = f(e) = \varepsilon$$
$$f(a') \circ f(a) = f(a' * a) = f(e) = \varepsilon$$

es decir, $f(a)$ tiene por simétrico a $f(a')$, o sea, es $f(a)' = f(a')$.

3. Obsérvese primeramente que $f(\bar{G}) \neq \emptyset$, ya que $\bar{G} \neq \emptyset$; sólo hay, pues, que probar que, si $\alpha, \beta \in f(\bar{G})$, entonces $\alpha \circ \beta' \in f(\bar{G})$. Así ocurre, puesto que

como $\alpha = f(a)$ y $\beta = f(b)$ para ciertos $a, b \in \bar{G}$ y $\bar{G}$ es un subgrupo de G, será:

$$\alpha, \beta \in f(\bar{G}) \Rightarrow a, b \in \bar{G} \Rightarrow a * b' \in \bar{G} \Rightarrow f(a * b') \in f(\bar{G}) \Rightarrow$$
$$\Rightarrow f(a) \circ f(b') \in f(\bar{G}) \Rightarrow f(a) \circ f(b)' \in f(\bar{G}) \Rightarrow \alpha \circ \beta' \in f(\bar{G})$$

4. Por ser $\bar{H} \neq \emptyset$, también es $f^{-1}(\bar{H}) \neq \emptyset$. Sólo queda, pues, por probar que si $a, b \in f^{-1}(\bar{H})$ entonces $a * b' \in f^{-1}(\bar{H})$. Así es, ya que recurriendo a $\alpha = f(a)$ y $\beta = f(b)$ y como $\alpha, \beta \in \bar{H}$ y $\bar{H}$ es subgrupo, resulta que:

$$a, b \in f^{-1}(\bar{H}) \Rightarrow \alpha, \beta \in \bar{H} \Rightarrow \alpha \circ \beta' \in \bar{H} \Rightarrow f(a) \circ f(b)' \in \bar{H} \Rightarrow$$
$$\Rightarrow f(a) \circ f(b') \in \bar{H} \Rightarrow f(a * b') \in \bar{H} \Rightarrow a * b' \in f^{-1}(\bar{H})$$

5. Para cualesquiera que sean $\alpha, \beta \in H$, como existen unos $a, b \in G$ tales que $\alpha = f(a)$ y $\beta = f(b)$ y por ser f isomorfismo. resulta que:

$$f^{-1}(\alpha \circ \beta) = f^{-1}(f(a) \circ f(b)) = f^{-1}(f(a * b)) = a * b = f^{-1}(\alpha) \circ f^{-1}(\beta)$$

luego f^{-1} es homomorfismo. Como f^{-1} también es biyectiva, es entonces un isomorfismo.

6. Para cualesquiera $a, b \in G$, como f y g son homomorfismo, se tiene:

$$(g \circ f)(a * b) = g(f(a * b)) = g(f(a) \triangle f(b)) = g(f(a)) \bigstar g(f(b)) =$$
$$= (g \circ f)(a) \bigstar (g \circ f)(b)$$

como había que comprobar.

NUCLEO E IMAGEN DE UN HOMOMORFISMO

[391]

Si $f: (G, *) \to (H, \circ)$ es un homomorfismo de grupos, se verifica que:

1. **El conjunto imagen $f(G)$ es un subgrupo de $(H, \circ)$. A este subgrupo se le llama *imagen* del homomorfismo f y se le denota poniendo Im (f).**
2. **Llamando ε al elemento neutro de H, el conjunto**

$$f^{-1}(\varepsilon) = \{a \in G / f(a) = \varepsilon\}$$

es un subgrupo de $(G, *)$. A este subgrupo se le llma *núcleo* del homomorfismo f y se le denota poniendo $N(f)$, Nuc (f) o Ker (f) (*).

(*) En inglés, «núcleo» se dice «Kernel»; en alemán, se dice «Kern».

DEMOSTRACION

1. Como G es un subgrupo de sí mismo, de la anterior propiedad 3 de [390] se desprende que $f(G)$ es un subgrupo de $(H, \circ)$.
2. Como $\{\varepsilon\}$ es un subgrupo de $(H, \circ)$, de la anterior propiedad 4 de [390] se deduce que $f^{-1}(\varepsilon)$ es un subgrupo de $(G, *)$.

EJEMPLO

La aplicación $f: \mathbb{Z} \to \mathbb{C}^*$ que a cada número entero n le hace corresponder el complejo $f(n) = i^n$ es un homomorfismo del grupo aditivo $(\mathbb{Z}, +)$ en el grupo multiplicativo $(\mathbb{C}^*, \cdot)$, ya que:

$$f(n+m) = i^{n+m} = i^n \cdot i^m = f(n) \cdot f(m)$$

La imagen de este homomorfismo esta formada por las potencias enteras de i, esto es, Im $(f) = \{1, -1, i, -i\}$.
El núcleo de este homomorfismo es $f^{-1}(1)$, esto es, el conjunto $\{n \in \mathbb{Z} / i^n = 1\}$; por ello Nuc $(f) = \{n = 4p / p \in \mathbb{Z}\}$.

HOMOMORFISMOS INYECTIVOS

[392]

Un homomorfismo de grupos es inyectivo si y sólo si su núcleo está formado solamente por el elemento neutro. Esto es, el homomorfismo de grupos $f: (G, *) \to (H, \circ)$ es inyectivo si y sólo si Nuc $(f) = \{e\}$ (e = elemento neutro de G).

DEMOSTRACION

Llamemos ε al elemento neutro de H; sabemos que $f(e) = \varepsilon$ y, por ello, que $e \in$ Nuc (f).

- Supongamos primero que f es inyectivo. Si $a \in$ Nuc (f), o sea, si $f(a) = \varepsilon$, como $f(e) = \varepsilon$ y f es inyectiva resulta que $a = e$, luego Nuc (f) sólo consta del elemento e, es decir, Nuc $(f) = \{e\}$.
- Supongamos ahora que Nuc $(f) = \{e\}$. Si $a, b \in G$ son tales que $f(a) = f(b)$, se habrán de verificar las siguientes relaciones (cada una de ellas es consecuencia de su anterior):

$$f(a) * f(b)' = \varepsilon, \quad f(a) * f(b') = \varepsilon, \quad f(a * b') = \varepsilon$$
$$a * b' \in \text{Nuc } (f), \quad a * b' = e, \quad a = b$$

luego f es inyectiva, como había que comprobar.

GRUPOS DE SUSTITUCIONES

Nos vamos a ocupar aquí de las biyecciones de un conjunto finito, de n elementos, en si mismo; éstas, que reciben el nombre de sustituciones, forman grupo respecto de la composición de aplicaciones. Este grupo no se trae aquí tanto como ejemplo, que lo es y bueno, como por la necesidad que de él se tiene al estudiar los determinantes.

[393]

Se llaman *sustituciones* de n elementos a las biyecciones del conjunto $N=\{1, 2, ..., n\}$ en sí mismo(*). Para representar a una sustitución $\sigma: N \to N$ se acostumbra a poner

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \downarrow & \downarrow & & \downarrow \\ h & k & \cdots & l \end{pmatrix} \quad \begin{pmatrix} \text{donde se han llamado} \\ h=\sigma(1),\ k=\sigma(2),\ ...,\ l=\sigma(n) \end{pmatrix}$$

La sustitución σ viene dada por la sucesión $(h, k, ..., l)$ en que se transforma la sucesión $(1, 2, ..., n)$ de partida.

El conjunto $\mathcal{S}_n$, de las $n!$ sustituciones de n elementos, es un grupo respecto de la composición ($\circ$) de aplicaciones, al que se llama *grupo simétrico* de orden n. Este grupo no es conmutativo (para $n>2$).

DEMOSTRACION

La composición ($\circ$) de dos biyecciones de N en N es otra biyección de N en N y, por ello, $\circ$ es operación interna para $\mathcal{S}_n$. Como la composición de aplicaciones es asociativa, $\circ$ es operación asociativa. La aplicación idéntica i es tal que $\sigma \circ i = i \circ \sigma = \sigma$ para toda $\sigma \in \mathcal{S}_n$, es decir, i es el elemento neutro de la estructura $(\mathcal{S}_n, \circ)$. Toda biyección $\sigma \in \mathcal{S}_n$ tiene biyección recíproca $\sigma^{-1} \in \mathcal{S}_n$, verificándose que $\sigma^{-1} \circ \sigma = \sigma \circ \sigma^{-1} = i$; esto es, todo elemento $\sigma \in \mathcal{S}_n$ tiene simétrico $\sigma^{-1} \in \mathcal{S}_n$. De lo dicho se desprende que $(\mathcal{S}_n, \circ)$ es un grupo.

Para comprobar que este grupo no es conmutativo, bastará con mostrar uno de los muchos casos en los que, para ciertos $\sigma, \tau \in \mathcal{S}_n$, es $\tau \circ \sigma \neq \sigma \circ \tau$; uno de tales casos es:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots \\ \downarrow & \downarrow & \downarrow & \\ 1 & 3 & 2 & \cdots \end{pmatrix} \quad \text{y} \quad \tau = \begin{pmatrix} 1 & 2 & 3 & \cdots \\ \downarrow & \downarrow & \downarrow & \\ 3 & 2 & 1 & \cdots \end{pmatrix}$$

(donde los elementos siguientes al 3 se transforman en sí mismos).

(*) El conjunto N se puede sustituir por cualquier otro conjunto $N' = \{a_1, a_2, ..., a_n\}$, de n elementos; la naturaleza de estos elementos carece aquí de relevancia.

TRASPOSICIONES

[394]

Una sustitución, de n elementos, se dice que es una *trasposición* si deja invariantes a $n-2$ elementos y sólo altera a los otros dos, transformando a cada uno de ellos en el otro. La trasposición τ tal que $\tau(i)=j$, $\tau(j)=i$ y $\tau(h)=h$ para $h\neq i,j$ (aquí se supone $1\leqslant i<j\leqslant n$), se denota poniendo:

$$\tau=\begin{pmatrix}1 & \cdots & i & \cdots & j & \cdots & n\\ \downarrow & & \downarrow & & \downarrow & & \downarrow\\ 1 & \cdots & j & \cdots & i & \cdots & n\end{pmatrix} \quad \text{o de modo abreviado} \quad \tau=\begin{pmatrix}i & j\\ \downarrow & \downarrow\\ j & i\end{pmatrix}$$

La inversa de una trasposición τ, cualquiera, es ella misma; o sea, $\tau^{-1}=\tau$.

Teorema. Toda sustitución se puede expresar como composición de trasposiciones.

DEMOSTRACION

Veamos un modo de expresar cualquiera sustitución dada σ en forma de producto de trasposiciones. Para ello, si es $\sigma^{-1}(1)=p$, empecemos poniendo:

$$\underbrace{\begin{pmatrix}1 & 2 & 3 & \cdots & p & \cdots & n\\ \downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow\\ h & k & l & \cdots & 1 & \cdots & m\end{pmatrix}}_{\sigma=\text{sustitución dada}}=\underbrace{\begin{pmatrix}1 & h\\ \downarrow & \downarrow\\ h & 1\end{pmatrix}}_{\tau_1}\circ\underbrace{\begin{pmatrix}1 & 2 & 3 & \cdots & p & \cdots & n\\ \downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow\\ 1 & k & l & \cdots & h & \cdots & m\end{pmatrix}}_{\sigma_1}$$

donde σ_1 deja invariante el primer elemento (el 1). Si es $\sigma_1^{-1}(2)=q$, nos va a interesar expresar que:

$$\sigma_1=\begin{pmatrix}1 & 2 & 3 & \cdots & q & \cdots & n\\ \downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow\\ 1 & k & l & \cdots & 2 & \cdots & m\end{pmatrix}=\underbrace{\begin{pmatrix}2 & k\\ \downarrow & \downarrow\\ k & 2\end{pmatrix}}_{\tau_2}\circ\underbrace{\begin{pmatrix}1 & 2 & 3 & \cdots & q & \cdots & n\\ \downarrow & \downarrow & \downarrow & & \downarrow & & \\ 1 & 2 & l & \cdots & k & \cdots & m\end{pmatrix}}_{\sigma_2}$$

donde σ_2 deja invariante a los dos primeros elementos (el 1 y el 2). Continuando del modo descrito, se obtienen sustituciones σ_1, σ_2, σ_3, ... en las que cada vez hay un nuevo elemento que permanece invariante; por tanto, σ_{n-2} es una trasposición y se dispone ya de la descomposición pedida:

$$\sigma=\tau_1\circ\tau_2\circ\cdots\circ\tau_{n-2}$$

Hay que notar que este proceso se puede acortar si, en alguna de sus etapas, la sustitución σ_i fuese la identidad; en tal caso, la descomposición quedaría $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_i$.

SIGNO DE UNA SUSTITUCION

[395]

Dada una sustitución σ, de n elementos, se dice que $\sigma(i)$ y $\sigma(j)$ presentan *inversión* si $\sigma(i) - \sigma(j)$ tiene distinto signo que $i - j$; sea $I(\sigma)$ el número total de inversiones que forman, dos a dos, los elementos $\sigma(1)$, $\sigma(2)$, ..., $\sigma(n)$. Se llama *signo* de σ a $\varepsilon(\sigma) = (-1)^{I(\sigma)}$; según que sea $\varepsilon(\sigma) = 1$ o $\varepsilon(\sigma) = -1$, se dice que σ es de clase par o de clase impar, respectivamente. Toda trasposición es una sustitución impar.

Para cualesquiera sustituciones σ y σ', de n elementos, se verifica que

$$\varepsilon(\sigma' \circ \sigma) = \varepsilon(\sigma') \cdot \varepsilon(\sigma)$$

El signo, considerado como aplicación $\varepsilon: \mathcal{S}_n \to \{1, -1\}$, es un homomorfismo sobreyectivo del grupo simétrico $(\mathcal{S}_n, \circ)$ en el grupo multiplicativo $(\{1, -1\}, \cdot)$.

DEMOSTRACION

Vamos a realizar esta demostración por etapas sucesivas:

1. Al número de inversiones $I(\sigma)$ y al signo $\varepsilon(\sigma)$, de una sustitución σ, muchas veces se los llamará número de inversiones y signo de la sucesión $(\sigma(1), \sigma(2), ..., \sigma(n))$. Si $\sigma(i) = p$ y $\sigma(j) = q$, los elementos $\sigma(i)$ y $\sigma(j)$ formarán inversión si p y q aparecen en la sucesión $(\sigma(1), \sigma(2), ..., \sigma(n))$ en distinto orden relativo del que tienen en la sucesión natural $(1, 2, ..., n)$.
2. Si en una sucesión $S = (h, ..., p, ..., q, ..., k)$, de los números $1, 2, ..., n$, se intercambian entre sí dos cualesquiera de sus elementos, por ejemplo, los p y q, la nueva sucesión $S' = (h, ..., q, ..., p, ..., k)$, que así se forma, es de distinta clase que la sucesión S.

   Empecemos estudiando el caso, particular, en el que entre p y q no hay ningún elemento de S; al intercambiar entre sí a los p y q, en las inversiones de S sólo se produce una alteración: si p y q forman inversión en S no la forman en S' y, recíprocamente, si no la forman en S la forman en S'. Por tanto, S y S' son de distinta clase.

   Consideremos ahora el caso general; sea c el número de elementos de S que hay entre los p y q. Para pasar de S a S' vamos a intercambiar p, con el elemento que le precede y a repetir el proceso c veces, hasta que p ocupe el lugar anterior a q; después intercambiamos p con q; después intercambiamos q con el elemento que le antecede y repetimos el proceso

c veces, hasta que q ocupe el lugar que, en S, ocupaba p. Hemos realizado, pues, un total de $2c+1$ intercambios de elementos consecutivos. Como, según se obtuvo antes, con cada uno de estos intercambios se produce un cambio de clase, a la postre se han producido $2c+1$ cambios de clase (de par a impar o de impar a par), es decir, cuando se obtiene, al final, a la sucesión S' se llega con un cambio de clase respecto de S.

3. Si σ es una sustitución y τ es una trasposición, entonces σ y $\tau\circ\sigma$ tienen distinto signo, es decir, $\varepsilon(\tau\circ\sigma)=-\varepsilon(\sigma)$. En efecto: La sucesión (1, 2, ..., n) se transforma mediante σ en otra que llamaremos S; mediante $\tau\circ\sigma$, se obtiene otra sucesión S' que difiere de S en que dos de sus elementos aparecen intercambiados. Según se obtuvo antes, S y S' tienen distinto signo. Como la signatura de S es la de σ y la signatura de S' es la de $\tau\circ\sigma$, resulta que σ y $\tau\circ\sigma$ tienen distinto signo.
4. Dos descomposiciones distintas, de una misma sustitución σ, como composición de trasposiciones, $\sigma=\tau_1\circ\tau_2\circ\cdots\circ\tau_h$ y $\sigma=\tau'_1\circ\tau'_2\circ\cdots\circ\tau'_k$, son tales que h y k tienen la misma paridad. Para comprobarlo, recurriremos a la propiedad anterior, $\varepsilon(\tau\circ\sigma)=-\varepsilon(\sigma)$. Teniendo en cuenta que la identidad i es una sustitución que, obviamente, tiene signo unidad, se obtiene que:

$$\begin{aligned}\varepsilon(\sigma)&=\varepsilon[\tau_1\circ(\tau_2\circ\tau_3\circ\cdots\circ\tau_n)]=-\varepsilon(\tau_2\circ\tau_3\circ\cdots\circ\tau_n)=\\&=-\varepsilon[\tau_2\circ(\tau_3\circ\cdots\circ\tau_n)]=(-1)^2\varepsilon(\tau_3\circ\cdots\circ\tau_n)=\cdots=\\&=(-1)^{h-1}\varepsilon(\tau_h)=(-1)^{h-1}\varepsilon(\tau_n\circ i)=(-1)^h\varepsilon(i)=(-1)^h\end{aligned}$$

Análogamente se obtiene que $\varepsilon(\sigma)=(-1)^k$; por tanto $(-1)^h=(-1)^k$, es decir, h y k tienen la misma paridad.

5. Para cualesquiera sustituciones σ y σ', de n elementos, se verifica que $\varepsilon(\sigma'\circ\sigma)=\varepsilon(\sigma')\cdot\varepsilon(\sigma)$. En efecto, expresando σ y σ' como composición de trasposiciones, $\sigma=\tau_1\circ\tau_2\circ\cdots\circ\tau_p$ y $\sigma'=\tau'_1\circ\tau'_2\circ\cdots\circ\tau'_q$, resulta que $\varepsilon(\sigma)=(-1)^p$ y $\varepsilon(\sigma')=(-1)^q$. Por otra parte, como se verifica que $\sigma'\circ\sigma=$ $=\tau'_1\circ\tau'_2\circ\cdots\circ\tau'_q\circ\tau_1\circ\tau_2\circ\cdots\circ\tau_p$, nos encontramos con que

$$\varepsilon(\sigma'\circ\sigma)=(-1)^{q+p}=(-1)^q\cdot(-1)^p=\varepsilon(\sigma')\cdot\varepsilon(\sigma)$$

APENDICE

9

Anillos

Los campos numéricos $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$, en los que hay dos operaciones (suma y producto), son los primeros anillos y cuerpos con los que, a no dudarlo, nos hemos encontrado. No son, sin embargo, los únicos; ya que los polinomios y las matrices, por ejemplo, son otros anillos que no nos son ajenos. Aquí, lo realmente novedoso es el hecho de disponer de dos operaciones, que satisfacen a ciertos axiomas que se han obtenido como reflejo, más o menos fiel, de las propiedades de que gozan las sumas y productos de números, matrices, polinomios. A estas dos operaciones, definidas sobre unos mismos elementos, se las exige que guarden cierta relación, que queda plasmada en el axioma de distributividad.

Para definir los anillos, se recurre a un reducido número de axiomas, con lo que se obtiene una estructura abstracta muy general, que da cobijo a muchas y variadas estructuras concretas, particulares, que se estudian, así, simultáneamente. Tiempo hay, después, de añadir nuevos axiomas y ceñirse a los casos de mayor interés, entre los que destacan los cuerpos.

ANILLOS

[396]

Un conjunto A y dos operaciones internas definidas en él, a las que llamaremos suma $(+)$ y producto $(\cdot)$, se dice que forman un *anillo* $(A; +, \cdot)$ si se verifican las siguientes propiedades (axiomas de anillo):

1.ª $(A, +)$ es un grupo abeliano; esto es:

- La suma $+$ es asociativa $[(a+b)+c=a+(b+c),\ \forall a, b, c \in A]$.
- Existe neutro $o \in A[a+o=o+a=a,\ \forall a \in A]$.
- Cada $a \in A$ tiene un opuesto $-a \in A[a+(-a)=(-a)+a=o]$.
- La suma $+$ es conmutativa $[a+b=b+a,\ \forall a, b \in A]$.

2.ª El producto es asociativo; esto es:

- $(a \cdot b) \cdot c = a \cdot (b \cdot c)$, para cualesquiera $a, b, c \in A$.

3.ª El producto es distributivo respecto de la suma; esto es:

- $\begin{cases} a \cdot (b+c) = a \cdot b + a \cdot c \\ (b+c) \cdot a = b \cdot a + c \cdot a \end{cases}, \quad \forall a, b, c \in A$

- Se dice que $(A; +, \cdot)$ es *anillo unitario* si, además de las 1.ª, 2.ª y 3.ª, se verifica que: existe un elemento $u \in A$ que es unidad para el producto $(a \cdot u = u \cdot a = a,\ \forall a \in A)$.
- Se dice que $(A; +, \cdot)$ es *anillo conmutativo* si, además de las 1.ª, 2.ª y 3.ª, se verifica que el producto es conmutativo $(a \cdot b = b \cdot a,\ \forall a, b \in A)$.

EJEMPLOS

1. Los conjuntos numéricos $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$, con la suma y el producto usuales, son anillos conmutativos y unitarios.
2. Al conjunto A de las funciones $f: C \to \mathbb{R}$, definidas en un cierto conjunto C y que toman valores reales, se le puede dar estructura de anillo unitario y conmutativo si se le dota de la suma $(+)$ y del producto $(\cdot)$ usuales de funciones, esto es, de las operaciones definidas, para cualesquiera $f, g \in A$, mediante:

$$(f+g)(x) = f(x) + g(x) \quad \text{y} \quad (f \cdot g)(x) = f(x) \cdot g(x) \quad , \quad \forall x \in C$$

3. Sea $(G, *)$ un grupo abeliano y considérese el conjunto A de los homomorfismos f de dicho grupo en sí mismo. Este conjunto A es un anillo respecto de las operaciones suma $(+)$ y composición $(\circ)$ definidas, para cualesquiera $f, g \in A$, por:

$$(f+g)(x) = f(x) + g(x) \quad \text{y} \quad (g \circ f)(x) = g(f(x)) \quad , \quad \forall x \in G$$

(se definen así dos operaciones internas, pues estas $f+g$ y $g \circ f$ son homomorfismos de G en G). Este anillo es unitario pero no es conmutativo (en general).

PRIMERAS PROPIEDADES

[397]

Sea $(A; +, \cdot)$ un anillo. Además de verificarse, para $(A, +)$, las propiedades de los grupos abelianos (véase [387]), se cumple que:

1. $a \cdot o = o \cdot a = o, \quad \forall a \in A$ ($o \in A$ elemento neutro).
2. $\begin{cases} a \cdot (-b) = (-a) \cdot b = -a \cdot b \\ (-a) \cdot (-b) = a \cdot b \end{cases} \quad \forall a, b \in A$ (regla de los signos)

DEMOSTRACION

1. $a \cdot a = a \cdot a \Rightarrow a \cdot a = a \cdot (a+o) = a \cdot a + a \cdot o \Rightarrow$
$\Rightarrow a \cdot a - a \cdot a = a \cdot o \Rightarrow o = a \cdot o$
$a \cdot a = a \cdot a \Rightarrow a \cdot a = (a+o) \cdot a = a \cdot a + o \cdot a \Rightarrow$
$\Rightarrow a \cdot a - a \cdot a = o \cdot a \Rightarrow o = o \cdot a$
2. $o \cdot b = o \Rightarrow (-a+a) \cdot b = o \Rightarrow (-a) \cdot b + a \cdot b = o \Rightarrow$
$\Rightarrow (-a) \cdot b = -a \cdot b$
$a \cdot o = o \Rightarrow a \cdot (-b+b) = o \Rightarrow a \cdot (-b) + a \cdot b = o \Rightarrow$
$\Rightarrow a \cdot (-b) = -ab$
$(-a) \cdot (-b) = -[a \cdot (-b)] = [-ab] = a \cdot b$

ELEMENTOS INVERTIBLES DE UN ANILLO

[398]

Sea $(A; +, \cdot)$ un anillo unitario; denotemos por $u \in A$ a su unidad. Se llaman elementos invertibles del anillo aquellos $a \in A$ que admiten inverso $a^{-1} \in A$ (recuérdese que a^{-1} es tal que $a \cdot a^{-1} = a^{-1} \cdot a = u$).

1. Si $a, b \in A$ son dos elementos invertibles del anillo, entonces también lo es $a \cdot b$ y su inverso es:
$$(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$$
2. El conjunto $A^* = \{a \in A / \exists a^{-1} \in A\}$ es un grupo respecto del producto $(\cdot)$, al que se llama grupo multiplicativo del anillo A.

DEMOSTRACION

1. $$\left.\begin{array}{l}(a \cdot b) \cdot (b^{-1}a^{-1}) = a \cdot (b \cdot b^{-1}) \cdot a^{-1} = a \cdot u \cdot a^{-1} = a \cdot a^{-1} = u \\ (b^{-1} \cdot a^{-1}) \cdot (a \cdot b) = b^{-1} \cdot (a^{-1} \cdot a) \cdot b = b^{-1} \cdot u \cdot b = b^{-1} \cdot b = u\end{array}\right\} \Rightarrow$$
$$\Rightarrow \quad b^{-1} \cdot a^{-1} = (a \cdot b)^{-1}$$

2. De acuerdo con el apartado anterior, el producto es operación interna para A^*. Como el producto es asociativo en A, lo es en $A^* \subset A$. El elemento u es de A^* (tiene inverso, que es él mismo) y es la unidad de A^*, pues lo es en A. Sólo queda comprobar que si $a \in A^*$, el elemento a^{-1} (que existe en A) pertenece a A^*, es decir, que a^{-1} tiene inverso en A, lo que evidente ya que $(a^{-1})^{-1} = a \in A^*$.

SUBANILLOS

[399]

Sea $(A; +, \cdot)$ un anillo y sea B un subconjunto de A. Se dice que B es un subanillo del anillo $(A; +, \cdot)$ si B es una parte estable para las operaciones $+$ y $\cdot$ y, además, la estructura $(B, +, \cdot)$ es un anillo. Esto es:

Definición de subanillo. Se dice que $B \subset A$ es un subanillo de $(A; +, \cdot)$ si:

1. B es un subgrupo del grupo $(B; +)$.
2. $(a, b \in B) \Rightarrow (a \cdot b \in B)$ (o sea, $\cdot$ es operación para B).

Caracterización de los subanillos. El conjunto $B \subset A$ es un subanillo de $(A; +, \cdot)$ si y sólo si se verifica que:

i) $B \neq \emptyset$.
ii) $(a, b \in B) \Rightarrow [a - b \in B \quad \text{y} \quad a \cdot b \in B]$.

DEMOSTRACION

Esta caracterización de los subanillos es una consecuencia obvia de la caracterización de los subgrupos, que se analizó en [388].

EJEMPLOS

1. Un anillo A, cualquiera, tiene como subanillos (entre otros posibles) al propio A y a $\{o\}$ (o es el elemento nulo). A estos subanillos se les llama triviales; a los demás subanillos se les llama propios.
2. El anillo $(Z; +, \cdot)$ admite como subanillo al conjunto de los múltiplos de p (entero dado). El anillo $\mathbb{Q}$ es subanillo del $\mathbb{R}$ y éste lo es del $\mathbb{C}$.
3. Los polinomios, con una variable (o indeterminada) x, forman un anillo respecto de la suma y del producto usuales. Los polinomios que sólo tienen términos con potencias pares de x forman un subanillo. Obsérvese que, en este ejemplo, el anillo es unitario pero no lo es el subanillo.

DIVISORES DE CERO

El hecho que más suele sorprender, a quienes se inician en el estudio de los anillos, es la existencia de los «divisores de cero». Se conoce con este nombre a aquellos elementos que siendo no nulos dan producto nulo. Estos elementos no existen en los campos numéricos, pero si los hay en algunos anillos. Así, por ejemplo, consideremos el anillo $(A; +, \cdot)$, donde $A = \mathbb{R}^2$ y las operaciones $+$ y $\cdot$ están dadas por:

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$
$$(x_1, x_2) \cdot (y_1, y_2) = (x_1 y_1, x_2\, y_2)$$

Este anillo (que lo es, como se comprueba fácilmente) tiene divisores de cero. En efecto, los pares (3, 0) y (0, 8) son ambos distintos del elemento nulo, que es el (0, 0) y, sin embargo, al multiplicarles se obtiene el elemento nulo:

$$(3, 0) \cdot (0, 8) = (3 \cdot 0, 0 \cdot 8) = (0, 0)$$

[400]

En un anillo $(A; +, \cdot)$, se dice que un elemento no nulo $a \in A$ es un *divisor de cero* si existe algún otro elemento no nulo $b \in A$ tal que alguno de los $a \cdot b$ o $b \cdot a$ es el elemento nulo. Esto es:

[a y b divisores de cero] $\Leftrightarrow$ [$a \neq o$, $b \neq o$ y $(a \cdot b = o$ o $b \cdot a = o)$]

Propiedades. En un anillo $(A; +, \cdot)$, dado $a \in A$, se verifica que:

1. (a es divisor de cero) $\Leftrightarrow$ (a no es simplificable para el producto).
2. (a es invertible) $\Rightarrow$ (a no es un divisor de cero).

DEMOSTRACION

1. Si a es divisor de cero, entonces existe $b \neq o$ tal que $a \cdot b = o$ o $b \cdot a = o$, que se puede poner $a \cdot b = a \cdot o$ o $b \cdot a = o \cdot a$; si a fuese simplificable, de esto último se deduciría $b = o$, que es falso, luego a no es simplificable.
   Si a no es simplificable, existen b_1, $b_2 \in A$ tales que $b_1 \neq b_2$ y es $a \cdot b_1 = a \cdot b_2$ o $b_1 \cdot a = b_2 \cdot a$, que se puede poner $a \cdot (b_1 - b_2) = o$ o $(b_1 - b_2) \cdot a = o$ y, como $b_1 - b_2 \neq o$, de aquí se desprende que a es divisor de cero.
2. Como a es invertible, existe a^{-1} tal que $a \cdot a^{-1} = a^{-1} \cdot a = u$ (unidad). Por tanto, de ser $a \cdot b = o$ o $b \cdot a = o$, se deduciría:

$$a \cdot b = o \Rightarrow a^{-1} \cdot a \cdot b = a^{-1} \cdot o \Rightarrow u \cdot b = o \Rightarrow b = o$$
$$b \cdot a = o \Rightarrow b \cdot a \cdot a^{-1} = o \cdot a^{-1} \Rightarrow b \cdot u = o \Rightarrow b = o$$

   por tanto, a no puede ser un divisor de cero.

ANILLOS DE INTEGRIDAD

[401] Los anillos que carecen de divisores de cero se llaman *anillos de integridad*. Se acostumbra a llamar *dominio de integridad* a un anillo de integridad que sea unitario y conmutativo.

LOS ANILLOS DE CLASES RESIDUALES

Sin entrar en las demostraciones, que se dejan como ejercicios, vamos a describir aquí el «anillo de las clases de restas módulo p», donde $p > 1$ es un número natural dado.

- Para el conjunto $\mathbb{Z}$, de los enteros, se llama relación de «congruencia módulo p» a la relación binaria

$$(m \text{ relacionado con } n) \Leftrightarrow (m - n \text{ es múltiplo de } p)$$

  Esta relación es (véase el ejemplo 2 de [368]) una equivalencia; sus clases son $\bar{0}$. $\bar{1}$, $\bar{2}$, ..., $\overline{p-1}$, donde se ha llamado clase $\bar{r}$ al conjunto:

$$\bar{r} = \{r, r \pm p, r \pm 2p, r \pm 3p, \ldots\}$$

  El conjunto cociente se denota por $\mathbb{Z}/p$, es decir:

$$\mathbb{Z}/p = \{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{p-1}\}$$

- La anterior equivalencia es compatible con la suma y con el producto de $\mathbb{Z}$:

por tanto, estas operaciones inducen una suma y un producto en el conjunto cociente $\mathbb{Z}/p$, que están dadas por:

$$\bar{m}+\bar{n}=\overline{m+n} \quad \text{y} \quad \bar{m}\cdot\bar{n}=\overline{m\cdot n}$$

- La estructura algebraica $(\mathbb{Z}/p;\ +,\ \cdot\ )$ es un anillo conmutativo y unitario. El elemento nulo es la clase $\bar{0}$; el elemento unidad es la clase $\bar{1}$. Para $p=6$, por ejemplo, las tablas de la suma y del producto del anillo $(\mathbb{Z}/6;\ +,\ \cdot\ )$ son:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{5}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{0}$	$\bar{2}$	$\bar{4}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{0}$	$\bar{3}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{2}$	$\bar{0}$	$\bar{4}$	$\bar{2}$
$\bar{5}$	$\bar{0}$	$\bar{5}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

- Si p no es primo, es decir, si $p=h\cdot k$ (con $h,\ k\neq p$) para ciertos naturales h y k, entonces en el anillo $\mathbb{Z}/p$ hay divisores de cero, pues las clases $\bar{h}$ y $\bar{k}$ lo son. Así, en $\mathbb{Z}/6$, las clases $\bar{2}$ y $\bar{3}$ son divisores de cero, pues $\bar{2}\cdot\bar{3}=\bar{3}\cdot\bar{2}=\bar{0}$. Si p es primo, entonces en $\mathbb{Z}/p$ no hay divisores de cero; este anillo es, ahora, un dominio de integridad.

HOMOMORFISMOS DE ANILLOS

Respondiendo a la misma idea de linealidad que condujo al concepto de homomorfismo de grupos, se obtienen los homomorfismos entre anillos. La linealidad, que en los grupos se exigió para una operación, se requerirá aquí para las dos: para la suma y para el producto:

[402]

Dados dos anillos $(A;\ +,\ \cdot\ )$ y $(B;\ +,\ \cdot\ )$(*), se dice que una aplicación $f:A\to B$ es un *homomorfismo* del anillo $(A;\ +,\ \cdot\ )$ en el anillo $(B;\ +,\ \cdot\ )$ si, para cualesquiera que sean los elementos $a,\ b\in A$, se verifica que:

$$f(a+b)=f(a)+f(b) \quad \text{y} \quad f(a\cdot b)=f(a)\cdot f(b)$$

El homomorfismo de anillos $f:(A;\ +,\ \cdot\ )\to(B;\ +,\ \cdot\ )$ se dice que es un *isomorfismo* si la aplicación $f:A\to B$ es una biyección. Un homomorfismo $\varphi:(A;\ +,\ \cdot\ )\to(A;\ +,\ \cdot\ )$, de un anillo en sí mismo, recibe el nombre de *endomorfismo*. Un endomorfismo $\varphi:(A;\ +,\ \cdot\ )\to(A;\ +,\ \cdot\ )$ se dice que es un *automorfismo* si la aplicación $\varphi:A\to A$ es biyectiva.

(*) Las operaciones de los anillos A y B serán en general distintas, en uno y otro, aunque se denoten de igual modo (ambas con $+$, para la primera y con $\cdot$ para la segunda operación).

PROPIEDADES

[403]

Si $f: A \to B$ es un homomorfismo del anillo $(A; +, \cdot)$ en el anillo $(B; +, \cdot)$ se verifica que:

1. $f(o) = o$ (donde o denota a los neutros, de A y de B).
2. $f(-a) = -f(a)$, $\forall\ a \in A$.
3. ($\bar{A}$ es subanillo de A) $\Rightarrow$ ($f(\bar{A})$ es subanillo de B).
4. ($\bar{B}$ es subanillo de B) $\Rightarrow$ ($f^{-1}(\bar{B})$ es subanillo de A).
5. (f es un isomorfismo) $\Rightarrow$ (f^{-1} es un isomorfismo).
6. Si $f: (A; +, \cdot) \to (B; +, \cdot)$ y $g: (B; +, \cdot) \to (C; +, \cdot)$ son homomorfismos de anillos, entonces también lo es $g \circ f: (A; +, \cdot) \to (C; +, \cdot)$.

DEMOSTRACION

Nótese que f es un homomorfismo del grupo $(A, +)$ en el grupo $(B, +)$; por ello, se verifican aquí todas las propiedades de los homomorfismos de grupos (véase [390]). En particular, las 1 y 2 ya están probadas, también sabemos que $f(\bar{A})$ es subgrupo de B y que $f^{-1}(\bar{B})$ es subgrupo de A; la propiedad 5 ya está probada y, en la 6, ya se sabe que $g \circ f$ es homomorfismo de grupos. Sólo hay, pues, que comprobar aquello que, por afectar al producto, no quedó demostrado al estudiar los homomorfismos de grupo. Pero aun estas cuestiones pendientes no precisan de demostración expresa, ya que se prueban sin más que sustituir, en lo que se dijo para grupos en [390], las operaciones generales $*$, $\circ$, etc., por el producto $(\cdot)$.

NUCLEO E IMAGEN DE UN HOMOMORFISMO

[404]

Sea $f: (A; +, \cdot) \to (B; +, \cdot)$ un homomorfismo de anillos. Se verifica:

- El conjunto imagen $f(A)$ es un subanillo de $(B; +, \cdot)$, al que se llama *imagen* del homomorfismo f y se le representa poniendo Im (f).
- Si el anillo A es unitario y su unidad es u, entonces Im (f) es anillo unitario y su unidad es $f(u)$.

- El conjunto $f^{-1}(o) = \{a \in A / f(a) = o\}$, donde $o \in B$ es el elemento neutro, es un subanillo de $(A; +, \cdot)$, al que se llama *núcleo* del homomorfismo f y se le representará por $N(f)$, Nuc (f) o Ker (f).
- Si $a \in A$ y si $n \in$ Nuc (f), entonces los elementos $a \cdot n$ y $n \cdot a$ también son del núcleo de f.

DEMOSTRACION

1. Como A es un subanillo de si mismo, de la anterior propiedad 3 se desprende que $f(A)$ es un subanillo de $(B; +, \cdot)$.

   Si A tiene unidad, u, entonces para cualquiera que sea $b \in f(A)$, como $b = f(a)$ para algún $a \in A$, se puede poner:

$$b \cdot f(u) = f(a) \cdot f(u) = f(a \cdot u) = f(a) = b$$

$$f(u) \cdot b = f(u) \cdot f(a) = f(u \cdot a) = f(a) = b$$

   es decir, $f(u)$ ha resultado ser la unidad de Im (f).

2. Como $\{o\}$ es un subanillo del anillo $(B; +, \cdot)$, de la anterior propiedad 4 se deduce que $f^{-1}(o)$ es un subanillo de $(A; +, \cdot)$.

   Si $a \in A$ y $n \in$ Nuc (f), entonces $a \cdot n$ y $n \cdot a$ pertenecen al núcleo, ya que:

$$f(a \cdot n) = f(a) \cdot f(n) = f(a) \cdot o = o$$

$$f(n \cdot a) = f(n) \cdot f(a) = o \cdot f(a) = o$$

EJEMPLO

Considérense los anillos $\mathbb{Z}$ (de los enteros) y $\mathbb{Z}/p$ (de las clases de restos módulos p). Sea $f: \mathbb{Z} \to \mathbb{Z}/p$ la aplicación definida por

$$f(r) = \bar{r} \quad , \quad \forall r \in \mathbb{Z}$$

Esta aplicación es un homomorfismo de anillos ya que

$$f(r+s) = \overline{r+s} = \bar{r} + \bar{s} = f(r) + f(s)$$

$$f(r \cdot s) = \overline{r \cdot s} = \bar{r} \cdot \bar{s} = f(r) \cdot f(s)$$

Como f es sobreyectiva, la imagen de f es $\text{Im}(f) = \mathbb{Z}/p$. El núcleo de f está formado por los múltiplos de p, que son los elementos de $\mathbb{Z}$ que se transforman en la clase $\bar{0}$.

IDEALES DE UN ANILLO

[405]

En un anillo $(A; +, \cdot)$, un conjunto no vacío $I \subset A$ se dice que es un *ideal* si se verifican las dos condiciones siguientes:

$$(i, j \in I) \Rightarrow i - j \in I$$

$$(i \in I,\ a \in A) \Rightarrow (a \cdot i,\ i \cdot a \in I)$$

OBSERVACIONES

- Los requisitos que se le exigen a I para ser un ideal son propiedades a las que, según se acaba de comprobar en [404], verifican los núcleos de los homomorfismos de anillos. El núcleo de un homomorfismo es, pues, un ideal del anillo origen.
- Nótese que, en un anillo, los ideales son subanillo. El recíproco es falso en general; así, $\mathbb{Z}$ es un subanillo de $\mathbb{Q}$, pero $\mathbb{Z}$ no es un ideal de $\mathbb{Q}$.

APENDICE

10

Cuerpos

En la definición de anillo se fue mucho más exigente con la suma que con el producto, al que bastaba con pedirle que fuera asociativo (además de la distributividad, que afecta a la suma y al producto). Al requerir del producto que, como ocurre en los campos numéricos $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$, sea conmutativo, sea unitario y, lo que es más importante, que todo elemento no nulo tenga inverso, se obtienen unos anillos particularmente importantes que se llaman cuerpos.

DEFINICION Y PROPIEDADES

[406]

Un anillo $(K; +, \cdot)$ se dice que es un *cuerpo* si $K^* = K - \{o\}$ (donde o es el elemento nulo) es un grupo abeliano respecto del producto $(\cdot)$. Los cuerpos son, pues, los anillos unitarios y conmutativos $(K; +, \cdot)$ en los que, además, todo $a \in K^*$ admite inverso $a^{-1} \in K^*$ (de manera que $a \cdot a^{-1} = u$ unidad de K).

OBSERVACIONES

1. Hay quienes llaman cuerpo a la estructura que resulta al prescindir, en la anterior definición, de la conmutatividad del producto; a una tal estructura la llamaremos nosotros cuerpo no conmutativo.
2. Todo cuerpo tiene, al menos, dos elementos: el elemento nulo de la suma (que hemos llamado o) y el elemento unidad del producto (que hemos llamado u).

PROPIEDADES

[407]

Para los cuerpos, además de todas las propiedades de los anillos, también se verifica que:

1. (K es un cuerpo) $\Rightarrow$ (K es un dominio de integridad).
2. Un cuerpo K sólo tiene dos ideales, que son K y $\{o\}$.
3. La ecuación $a \cdot x = b$ (donde $a, b \in K$ son datos, $a \neq o$, y $x \in K$ es la incógnita) tiene solución única en el cuerpo K, que es la $x = a^{-1} \cdot b$.

DEMOSTRACION

1. Como los elementos que tienen inverso no son divisores de cero, en un cuerpo K no hay divisores de cero y, al ser además unitario y conmutativo, K es entonces dominio de integridad.
2. Si $I \neq \{o\}$ es un ideal de K, hay que probar que $I = K$. Como $I \neq \{o\}$, hay algún $i \in I$ con $i \neq o$; cualquier $a \in K$ se puede poner $a = i \cdot (i^{-1} \cdot a)$ y, como $i \in I$, como $i^{-1} \cdot a \in K$ y como I es un ideal, se concluye que $a = i \cdot (i^{-1} \cdot a)$ es de I. Resulta, pues, que $a \in K \Rightarrow a \in I$, es decir $I = K$.
3. Cada una de las siguientes relaciones es equivalente a la que va tras ella:

$$a \cdot x = b, \quad a^{-1} \cdot (a \cdot x) = a^{-1} \cdot b, \quad (a^{-1} \cdot a) \cdot x = a^{-1} \cdot b, \quad x = a^{-1} \cdot b$$

SUBCUERPO

[408]

Sea $(K; +, \cdot)$ un cuerpo y sea H un subconjunto de K. Se dice que H es un *subcuerpo* de $(K; +, \cdot)$ si H es una parte estable para las operaciones $+$ y $\cdot$ y tal que la estructura $(H; +, \cdot)$ es un cuerpo.

El conjunto $H \subset K$ es un subcuerpo del cuerpo $(K; +, \cdot)$ si y sólo si:

1. H es un subanillo de $K; +, \cdot)$; es decir

$$[a, b \in H] \Rightarrow [a - b \in H \text{ y } a \cdot b \in H]$$

2. $[a \in H, \ a \neq o] \Rightarrow a^{-1} \in H$.

DEMOSTRACION

Supongamos primero que se cumplen las condiciones 1 y 2 y veamos que, entonces, H es un subcuerpo. Hay que cerciorarse de dos cosas:

- Que $(H; +, \cdot)$ es un anillo; esto es así, según se dice en 1.
- Que $(H^*, \cdot)$ es grupo; esto, de acuerdo con la caracterización de los subgrupos (véase [388]), es consecuencia de 2 y de que $H^* \neq \emptyset$ (se ha llamado $H^* = H - \{o\}$).

El recíproco es aún más obvio; no precisa de comentarios.

EJEMPLOS

El cuerpo $\mathbb{Q}$ es un subcuerpo de $\mathbb{R}$; el cuerpo $\mathbb{R}$ es subcuerpo de $\mathbb{C}$. Los números reales del tipo $a + b\sqrt{2}$, para $a, b \in \mathbb{Q}$, forman un subcuerpo de $\mathbb{R}$.

CARACTERISTICA DE UN CUERPO

[409]

Sea $(K; +, \cdot)$ un cuerpo; sea $o \in K$ el elemento nulo del cuerpo; sea $a \neq o$ un elemento del cuerpo K. Para cada $n \in \mathbb{N}$, se conviene en poner:

$$na = a + a + \overset{n}{\dots} + a$$

- Si existe $p \in \mathbb{N}$ tal que $pa = o$, se dice que K tiene característica finita y se llama característica de K al menor $p \in \mathbb{N}$ tal que $pa = o$. Esta definición no depende del elemento $a \neq o$ que se tome de K; generalmente se toma $a = u$ (unidad de K).
- Si $na \neq o$ para todo n, se dice que K tiene característica infinita (esta definición tampoco depende del elemento $a \neq o$ que se tome en K).

DEMOSTRACION

En el caso de característica finita, hay que comprobar que si $p \in \mathbb{N}$ es el menor número natural tal que $pu = o$, entonces se verifican dos cosas: 1.ª) que $pa = o$; y 2.ª) que para $q < p$ es $qa \neq o$. Así ocurre:

1.ª $pa = a + a + \overset{p}{\dots} + a = u \cdot a + u \cdot a + \overset{p}{\dots} + u \cdot a = (u + u + \overset{p}{\dots} + u) \cdot a =$
$= (pu) \cdot a = o \cdot a = o$

2.ª Si para algún $q < p$ fuese $qa = o$, resultaría que

$$o = qa = q(u \cdot a) = u \cdot a + u \cdot a + \overset{q}{\dots} + u \cdot a = (u + u + \overset{q}{\dots} + u) \cdot a = (qu) \cdot a$$

Como es $q < p$, ha de ser $qu \neq o$. Al ser $a \neq o$ y $qu \neq o$, no puede ser $(qu) \cdot a = o$, pues en un cuerpo no hay divisores de cero, es decir, se ha llegado a una contradicción que obliga a rechazar que sea $qa = o$ para algún $q \in \mathbb{N}$ menor que p.

Para el caso de característica infinita, hay que probar que, si $nu \neq o$ para todo $n \in \mathbb{N}$ y dado $a \neq o$ en K, se verifica que $na \neq o$ para todo $n \in \mathbb{N}$. Así ocurre, ya que si fuese $na = o$ para algún $n \in \mathbb{N}$, sería

$$o = na = n(u \cdot a) = u \cdot a + u \cdot a + \overset{n}{\dots} + u \cdot a = (u + u + \overset{n}{\dots} + u) \cdot a = (nu) \cdot a$$

lo cual no es posible ya que, como $nu \neq o$ y $a \neq o$ y dado que en los cuerpos no hay divisores de cero, ha de ser $(nu) \cdot a \neq o$.

HOMOMORFISMOS DE CUERPOS

Los homomorfismos entre cuerpos son los ya conocidos homomorfismos entre anillos, en el caso particular de que éstos sean cuerpos. No hay, pues, ahora una nueva definición de homomorfismo; lo que si hay son propiedades específicas de estos homomorfismos particulares.

[410]

Dados dos cuerpos $(K; +, \cdot)$ y $(H; +, \cdot)$, se dice que una aplicación $f: K \to H$ es un homomorfismo del primer cuerpo en el segundo si, para cualesquiera $a, b \in K$, se verifica que:

$$f(a+b) = f(a) + f(b) \quad \text{y} \quad f(a \cdot b) = f(a) \cdot f(b)$$

Dicho homomorfismo es: 1) homomorfismo del grupo $(K, +)$ en el grupo $(H; +)$; 2) un homomorfismo del grupo $(K^*, \cdot)$ en el grupo $(H^*, \cdot)$ (*); y 3) un homomorfismo del anillo $(K; +, \cdot)$ en el anillo $(H; +, \cdot)$.

(*) $K^* = K - \{o\}$ y $H^* = H - \{o\}$; aquí, o denota el elemento neutro, tanto de K como de H.

OBSERVACIONES

Para los homomorfismos de cuerpos, como son homomorfismo de anillos y, también, homomorfismos entre los respectivos grupos multiplicativos, se verifican ahora todas las propiedades de estos homomorfismo (véanse [390] y [403]. Así, por ejemplo, si $f: K \to H$ es un homomorfismo de cuerpos, se verifica que $f(o) = o, f(-a) = -f(a)$ para todo $a \in K, f(a^{-1}) = f(a)^{-1}$ para todo $a \in K^*$. Además de estas propiedades, hay otras, específicas, de las que pasamos a ocuparnos.

PROPIEDADES

[411]

Si $f: K \to H$ es un homomorfismo de un cuerpo $(K; +, \cdot)$ en otro cuerpo $(H; +, \cdot)$, se verifica que:

1. La imagen de f es un subcuerpo de H.
2. El núcleo de f es K o es $\{o\}$.
3. f es inyectivo, salvo si es el homomorfismo nulo.

DEMOSTRACION

1. Como f es un homomorfismo de anillos, sabemos que Im (f) es un subanillo de H. Sólo nos queda probar, pues, que dado $b \in$ Im (f), $b \neq o$, se verifica que $b^{-1} \in$ Im (f). Así es, ya que: como existe $a \in K$ tal que $f(a) = b$, $a \neq o$, y por ser K cuerpo, se sabe que $a^{-1} \in K$ y, por ello, $f(a^{-1}) \in$ Im (f), es decir, $f(a)^{-1} \in$ Im (f), que es $b \in$ Im (f).
2. Nuc (f) es el núcleo de un homomorfismo de anillos y, por ello, es un ideal de K. Como, por otra parte, los únicos ideales de un cuerpo K son K y $\{o\}$, la propiedad queda ya probada.
3. De acuerdo con la propiedad anterior, sólo puede ocurrir que sea Nuc $(f) = K$ o que sea Nuc $(f) = \{o\}$. Si Nuc $(f) = K$, resulta que $f(a) = o$ para todo $a \in K$, es decir, f es el homomorfismo nulo. Si Nuc $(f) = \{o\}$, según sabemos de los homomorfismos de anillo, entonces f es inyectivo.

APENDICE

11

Los ángulos y su medida

Aquí se pretende fundamentar el concepto de ángulo, yendo más allá de las ideas elementales, pero nada despreciables, que sobre este particular se manejan en la enseñanza preuniversitaria. Bueno es recordar que la noción elemental de ángulo dice, poco más o menos, que éste es la figura geométrica constituida por dos semirrectas que tienen el mismo origen.

ROTACIONES BIDIMENSIONALES

Recuérdese (véase [148] a [151]) que, en un espacio vectorial euclídeo V, las transformaciones ortogonales son los automorfismos f que conservan el producto escalar. En bases ortonormales, una matriz A está asociada a una transformación ortogonal si y sólo si $A^tA = I$, o sea, si A es ortogonal. También sabemos que $\det f$ vale 1 ó -1; en el primer caso se dice que f es una rotación. Las rotaciones en V forman grupo, respecto de la composición de aplicaciones, al que se representa por $\mathbb{O}^+(V)$. En bases ortonormales, las matrices de las rotaciones tienen determinante valiendo 1 y se llaman matrices ortogonales directas o de rotación; ellas forman grupo multiplicativo, que se denota por $\mathbb{O}^+(n)$, donde $n = \dim V$.

[412]

Sea V_2 un espacio vectorial euclídeo de dimensión 2; $\mathbb{O}^+(V_2)$ denota al grupo de las rotaciones en V_2 y $\mathbb{O}^+(2)$ representa al grupo (multiplicativo) de las matrices de rotación de tamaño 2×2. Se verifica que:

1. Las matrices del grupo $\mathbb{O}^+(2)$ son las del tipo
$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{con } a^2 + b^2 = 1 \quad (a, b \in \mathbb{R})$$
2. El grupo $\mathbb{O}^+(2)$ es isomorfo al grupo multiplicativo $\mathcal{U}$ de los números complejos de módulo unidad ($A \longmapsto a + bi$ es un isomorfismo).
3. Los grupos $\mathbb{O}^+(2)$ y $\mathbb{O}^+(V_2)$ son conmutativos.
4. Si $\vec{u}, \vec{v} \in V_2$ son vectores unitarios, existe una sola rotación $f \in \mathbb{O}^+(V_2)$ tal que $f(\vec{u}) = \vec{v}$.
5. Si V_2 está orientado(*), dada una rotación $f \in \mathbb{O}^+(V_2)$, se verifica que las matrices $A \in \mathbb{O}^+(2)$ asociadas a f en las distintas bases ortonormales con orientación positiva de V_2 son todas iguales (son la misma).

(*) Atendiendo a su orientación, las bases de V_2 se dividen en dos clases (véase [095]). Se considera como positiva la orientación de todas las bases de una de estas dos clases. Si $V_2 = \mathbb{R}^2$, la base canónica es la que se acostumbra a tomar como positivamente orientada.

DEMOSTRACION

1. Sea A una matriz de tamaño 2×2 cualquiera, como la que abajo se indica; para que A sea matriz de rotación han de verificarse las relaciones que se señalan:

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad \begin{matrix} a^2 + b^2 = 1 \text{ y } a^2 + c^2 = 1^{(1)} \\ a^2 + b^2 = 1 \text{ y } b^2 + d^2 = 1^{(2)} \end{matrix} \left| \begin{matrix} ac + bd = 0^{(3)} \\ ad - bc = 1^{(4)} \end{matrix} \right.$$

De (1) y (2) se deduce que $b^2 = c^2$ y $d^2 = a^2$, es decir, $c = \varepsilon b$ y $d = \varepsilon' a$, donde $\varepsilon = \pm 1$ y $\varepsilon' = \pm 1$. Llevando estas expresiones a (3) se obtiene que $\varepsilon = -\varepsilon'$ y de (4) se obtiene que $\varepsilon' = 1$. Así, pues, ha de ser

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{donde} \quad a^2 + b^2 = 1$$

2. La aplicación $F: \mathbb{O}^+(2) \to \mathcal{U}$, $A \mapsto F(A) = a + bi$ (donde A es la anterior última matriz) es obviamente una biyección; además es lineal respecto del producto, ya que:

$$F\left(\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}\right) = F\left(\begin{bmatrix} a\alpha - b\beta & -a\beta - b\alpha \\ a\beta + b\alpha & a\alpha - b\beta \end{bmatrix}\right) =$$

$$= (a\alpha - b\beta) + (a\beta + b\alpha)i$$

$$F\left(\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\right) F\left(\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}\right) = (a + bi)(\alpha + \beta i) = (a\alpha - b\beta) + (a\beta + b\alpha)i$$

3. El grupo $\mathbb{O}^+(2)$ es conmutativo ya que lo es el grupo $\mathcal{U}$, de los números complejos unitarios, que es isomorfo a aquél, como se acaba de comprobar. Como $\mathbb{O}^+(V_2)$ es isomorfo a $\mathbb{O}^+(2)$, la conmutatividad de éste conduce a la de aquél.
4. Sabemos que existe alguna base ortonormal $(\vec{e}_1, \vec{e}_2)$ de V_2 en la que $\vec{e}_1 = \vec{u}$. Sea (a, b) el par de coordenadas de $\vec{v}$ en dicha base, $\vec{v} = a\vec{e}_1 + b\vec{e}_2$; como $\vec{v}$ es unitario, será $a^2 + b^2 = 1$. Si existiera la rotación en la que $\vec{v} = f(\vec{u})$, su matriz sería

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

ya que su primera columna es la de las coordenadas de $f(\vec{e}_1) = \vec{v}$ y, además, se trata de una matriz de rotación. La rotación que en la base $(\vec{e}_1, \vec{e}_2)$ tiene a A por matriz es, pues, la única solución del problema.
5. Sean A y A' las matrices de f en dos bases ortonormales positivamente orientadas de f; si es Q la matriz del correspondiente cambio de coordenadas, sabemos que $A' = Q^{-1}AQ$. Como ambas bases son ortonormales y con la misma orientación, resulta que Q es ortogonal directa, esto es, Q es

matriz de una rotación, $Q \in \mathbb{O}^+(2)$. Resulta entonces que A y Q son matrices de $\mathbb{O}^+(2)$. Como este grupo es conmutativo, se tiene:

$$A' = Q^{-1}AQ = Q^{-1}(AQ) = Q^{-1}(QA) = (Q^{-1}Q)A = IA = A$$

como había que comprobar.

EJERCICIO

Sea V_2 un espacio vectorial euclídeo de dimensión 2, en el que se considera una base ortonormal. Hallar las matrices, en dicha base, de las rotaciones $f \in \mathbb{O}^+(V_2)$ tales que:

1. $f \circ f = i$
2. $f \circ f = -i$ (i = identidad en V_2)

RESOLUCION

Si es A la matriz pedida, debe verificarse que $A^2 = I$, en el caso primero, y que $A^2 = -I$, en el caso segundo. Recurriendo a la expresión general de una matriz de rotación, ha de ser:

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad A^2 = \begin{bmatrix} a^2 - b^2 & -2ab \\ 2ab & a^2 - b^2 \end{bmatrix} = \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

1. La ecuación $f \circ f = i$ equivale a $A^2 = I$, o bien, a:

$$\left.\begin{aligned} a^2 - b^2 &= 1 \\ 2ab &= 0 \end{aligned}\right\} \text{ cuyas soluciones son } \begin{cases} a = 1 & \text{y} \quad b = 0 \\ a = -1 & \text{y} \quad b = 0 \end{cases}$$

   Es decir, las soluciones son $A = I$ (f = identidad) y $A = -I$; esta última rotación se llama «simetría respecto del origen».
2. La ecuación $f \circ f = -i$ equivale a la $A^2 = -I$, o bien a:

$$\left.\begin{aligned} a^2 - b^2 &= -1 \\ 2ab &= 0 \end{aligned}\right\} \text{ cuyas soluciones son } \begin{cases} a = 0 & \text{y} \quad b = 1 \\ a = 0 & \text{y} \quad b = -1 \end{cases}$$

   Es decir, las soluciones son las dos matrices

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

ANGULOS

En la siguiente definición de ángulo vamos a echar mano de las *semirrectas vectoriales*. En un espacio vectorial V, cada vector no nulo $\bar{u} \in V$ engendra una semirrecta vectorial, que es el conjunto de todos los vectores $\lambda\bar{u}$, donde λ es cualquier escalar positivo ($\lambda > 0$). Los vectores de una semirrecta tienen, todos, la misma dirección y el mismo sentido.

[413]

Dado un par cualquiera de semirrectas (R, R') de un espacio vectorial euclídeo V_2 (de dimensión 2), existe una única rotación $f \in \mathbb{O}^+(V_2)$ tal que $f(R) = R'$; a esta rotación la representaremos por rot (R, R'). La relación $\simeq$ definida entre los pares de semirrectas de V_2 mediante

$$(R, R') \simeq (S, S') \Leftrightarrow \text{rot}\,(R, R') = \text{rot}\,(S, S')$$

es una equivalencia. Las clases de esta equivalencia se llaman *ángulos* (de semirrectas); la clase α de representante (R, R') se denota poniendo $\alpha = \text{áng}\,(R, R')$. Al conjunto de los ángulos le llamamos $\mathcal{A}$. Dados un ángulo $\alpha \in \mathcal{A}$ y una semirrecta R, existe una única semirrecta R' tal que $\alpha = \text{áng}\,(R, R')$. La *suma de ángulos* se define mediante:

$$\text{áng}\,(R, R') + \text{áng}\,(R', R'') = \text{áng}\,(R, R'') \quad \text{(relación de Chasles)}$$

(el resultado de la suma es independiente de la semirrecta R que se utilice para obtenerla). Con esta suma, $\mathcal{A}$ es un grupo conmutativo isomorfo al grupo $\mathbb{O}^+(V_2)$ de las rotaciones; la correspondencia que a cada ángulo $\alpha = \text{áng}\,(R, R')$ le hace corresponder la rotación $f = \text{rot}\,(R, R')$ es un isomorfismo de grupos. Este isomorfismo y su recíproco se denotan por «rot» y «áng»; es decir, las imágenes de $\alpha \in \mathcal{A}$ y de $f \in \mathbb{O}^+(V_2)$ se denotan por $\alpha \mapsto f = \text{rot}\,(\alpha)$ y por $f \mapsto \alpha = \text{áng}\,(f)$.

DEMOSTRACION

Sin entrar en cuestiones de detalle, las anteriores afirmaciones se pueden comprobar como sigue:

- Dadas R y R', hay que probar que existe $f \in \mathbb{O}^+(V_2)$ tal que $f(R) = R'$. Llamando $\vec{u}$ y $\vec{u}'$ a los vectores unitarios de las direcciones y sentidos de R y R', según vimos recientemente ([412], 4) existe una única rotación f tal que $f(\vec{u}) = \vec{u}'$; esta rotación es la buscada (que es única).
- La relación $\simeq$ es una equivalencia pues es, obviamente, reflexiva, simétrica y transitiva.
- Dado un ángulo $\alpha = \text{áng}\,(S, S')$ y una semirrecta R, si es $f = \text{rot}\,(S, S')$, la recta $R' = f(R)$ es tal que $\alpha = \text{áng}\,(R, R')$.
- Con la relación de Chasles siempre existe resultado para la suma. En efecto: dados $\alpha, \alpha' \in \mathcal{A}$ y eligiendo arbitrariamente una semirrecta R, existe R' tal que $\alpha = \text{áng}\,(R, R')$ y, disponiendo ya de la semirrecta R', existe R'' tal que $\alpha' = \text{áng}\,(R', R'')$; por tanto existe $\alpha + \alpha' = \text{áng}\,(R, R'')$.
- El anterior resultado $\alpha + \alpha'$ es independiente de la semirrecta R que se elija. En efecto: tomando (en lugar de R) otra semirrecta S, existen S' y S'' tales que $\alpha = \text{áng}\,(S, S')$ y $\alpha' = \text{áng}\,(S', S'')$. Por tanto:

$$\text{áng}\,(R, R') = \alpha = \text{áng}\,(S, S') \Rightarrow \text{rot}\,(R, R') = \text{rot}\,(S, S') = f$$
$$\text{áng}\,(R', R'') = \alpha' = \text{áng}\,(R', R'') \Rightarrow \text{rot}\,(R', R'') = \text{rot}\,(S', S'') = f'$$

Nótese, por otra parte, que dadas tres semirrectas T, T' y T'', la composición de las rotaciones g y g' determinadas por $g(T) = T'$ y $g'(T') = T''$, que es la rotación que transforma T en T'', también es la composición $g' \circ g$. Por ello,

$$\text{rot}\,(R, R'') = f' \circ f = \text{rot}\,(S, S''), \quad \text{luego} \quad \text{áng}\,(R, R'') = \text{áng}\,(S, S'')$$

- La correspondencia de $\mathcal{A}$ en $\mathbb{O}^+(V_2)$ que a cada ángulo $\alpha = \text{áng}\,(R, R')$ le asocia la rotación $f = \text{rot}\,(R, R')$ es obviamente una aplicación (no depende de la semirrecta R) biyectiva. Además, si en esta biyección $\alpha \rightarrowtail f$ y $\alpha' \rightarrowtail f'$, entonces $\alpha + \alpha' \rightarrowtail f' \circ f$ (según acabamos de obtener en la demostración anterior). Es decir, esta biyección que a cada ángulo le asocia «su» rotación, es un isomorfismo del grupo aditivo $\mathcal{A}$ en el grupo de las rotaciones $\mathbb{O}^+(V_2)$.

EJERCICIOS

1. Para cualesquiera que sean las semirrectas R, R', S y S', de un espacio vectorial euclídeo bidimensional V_2, pruébese que

$$\text{áng}\,(R, R') = \text{áng}\,(S, S') \quad \Leftrightarrow \quad \text{áng}\,(R, S) = \text{áng}\,(R', S')$$

2. Para cualesquiera que sean las semirrectas R y R', de un espacio vectorial euclídeo bidimensional V_2, y para cualquiera que sea la rotación $f \in \mathbb{O}^+(V_2)$, compruébese que

$$\text{áng}\,(R, R') = \text{áng}\,(f(R), f(R'))$$

RESOLUCION

1. Para abreviar, en lugar de áng (T, T'), pondremos $(\widehat{T, T'})$

$$(\widehat{R, R'}) = (\widehat{S, S'}) \quad \Leftrightarrow \quad (\widehat{R, R'}) + (\widehat{R', S}) = (\widehat{S, S'}) + (\widehat{R', S}) \quad \Leftrightarrow$$

$$\Leftrightarrow \quad (\widehat{R, S}) = (\widehat{R', S}) + (\widehat{S, S'}) \quad \Leftrightarrow \quad (\widehat{R, S}) = (\widehat{R', S'})$$

2. Sea $g \in \mathbb{O}^+(V_2)$ la rotación que transforma R en R'. Como el grupo $\mathbb{O}^+(V_2)$ es conmutativo, se verifica que $g \circ f = f \circ g$ y, por ello

$$g(f(R)) = f(g(R)) = f(R')$$

es decir, g es la rotación que transforma $f(R)$ en $f(R')$. Resulta, pues, que:

$$g : R \rightarrowtail g(R) = R' \quad \text{y} \quad g : f(R) \rightarrowtail g(f(R)) = f(R')$$

esto es, que g es la rotación del ángulo áng (R, R') y g es también la rotación del ángulo áng $(f(R), f(R'))$, luego estos dos ángulos son iguales, como había que comprobar.

[414]

Se llama *ángulo de dos vectores* $\bar{u}$ y $\bar{v}$, no nulos, de un espacio vectorial euclídeo V_2, al ángulo de las semirrectas que ellos engendran; esto es:

$$\text{áng}\,(\bar{u}, \bar{v}) = \text{áng}\,(R, S) \quad \text{donde} \quad \begin{cases} R = \{\lambda\bar{u}/\lambda > 0,\ \lambda \in \mathbb{R}\} \\ S = \{\lambda\bar{v}/\lambda > 0,\ \lambda \in \mathbb{R}\} \end{cases}$$

Si $\bar{u}$ y $\bar{v}$ son unitarios, entonces $\alpha = \text{áng}\,(\bar{u}, \bar{v})$ es el ángulo de la rotación $f \in \mathbb{O}^+(V_2)$ tal que $f(\bar{u}) = \bar{v}$.

- Se llama *ángulo nulo* (ω) al ángulo asociado a la rotación nula $f = o$; es decir, $\omega = \text{áng}\,(R, R) = \text{áng}\,(\bar{u}, \bar{u})$, para cualesquiera que sean la semirrecta R o el vector no nulo $\bar{u}$.
- Se llama *ángulo llano* ($\bar{\omega}$) al ángulo asociado a la rotación f tal que $f \circ f = i$ (identidad) (Ejercicio 1, [412]); es decir, $\bar{\omega} = \text{áng}\,(R, -R) = \text{áng}\,(\bar{u}, -\bar{u})$, para cualesquiera semirrecta R o vector no nulo $\bar{u}$.
- Se llaman *ángulos rectos* ($\pm\delta$) a los ángulos asociados a las rotaciones f tales que $f \circ f = -i$ (Ejercicio 2, [412]); es decir, $\pm\delta = \text{áng}\,(R, R^\perp) = \text{áng}\,(\bar{u}, \bar{u}^\perp)$, para cualquier semirrecta R o cualquier vector no nulo $\bar{u}$ y donde $R^\perp$ es a una de las dos semirrectas ortogonales a R y $\bar{u}^\perp$ es a uno de los dos vectores ortogonales a $\bar{u}$ de su mismo módulo.

SENO Y COSENO DE UN ANGULO

Sea V_2 un espacio vectorial euclídeo de dimensión 2. Vamos a introducir las «razones trigonométricas» de los ángulos en V_2.

[415]

1. Dada una rotación $f \in \mathbb{O}^+(V_2)$ y si $\bar{u} \in V_2$ es unitario, el número $\bar{u} \cdot f(\bar{u})$ es independiente del vector $\bar{u}$. Este número $\bar{u} \cdot f(\bar{u})$ se llama coseno de la rotación f o *coseno del ángulo* α de la rotación f.
2. Dada una rotación $f \in \mathbb{O}^+(V_2)$ y si $(\bar{u}_1, \bar{u}_2)$ es una base ortonormal de V_2, el número $\bar{u}_1 \cdot f(\bar{u}_2)$ es el mismo para todas las bases ortonormales $(\bar{u}_1, \bar{u}_2)$ con igual orientación. Si V_2 se considera orientado, el número $\bar{u}_1 \cdot f(\bar{u}_2)$, para cualquier base ortonormal positivamente orientada, se llama seno de la rotación f o *seno del ángulo* α de la rotación f.
3. Si V_2 está orientado entonces la matriz, A, de una rotación $f \in \mathbb{O}^+(V_2)$ es la misma (véase [412], 1 y 5) en todas las bases ortonormales positivamente orientadas y es del tipo:

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{para ciertas } a,\ b \in \mathbb{R} \quad \text{tales que } a^2 + b^2 = 1.$$

Pues bien, si α es el ángulo de la rotación f, se verifica que

$$\begin{matrix} \cos\alpha = a \\ \operatorname{sen}\alpha = b \end{matrix} \left(\begin{matrix} \text{obsérvese que} \\ \cos^2\alpha + \operatorname{sen}^2\alpha = 1 \end{matrix} \right) \quad \text{y} \quad A = \begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$$

DEMOSTRACION

1. Sea $\bar{v} \in V_2$ otro vector unitario cualquiera; hay que comprobar que $\bar{v} \cdot f(\bar{v})$ es igual a $\bar{u} \cdot f(\bar{u})$. Si es $g \in \mathbb{O}^+(V_2)$ la rotación tal que $\bar{v} = g(\bar{u})$ (véase [412], 4), como las rotaciones conmutan ($f \circ g = g \circ f$), se tiene:

$$\bar{v} \cdot f(\bar{v}) = g(\bar{u}) \cdot f(g(\bar{u})) = g(\bar{u}) \cdot g(f(\bar{u})) = \bar{u} \cdot f(\bar{u})$$

ya que g, por ser transformación ortogonal, conserva el producto escalar.

2. Sea $(\bar{v}_1, \bar{v}_2)$ otra base ortonormal igual orientada que la $(\bar{u}_1, \bar{u}_2)$; hay que comprobar que $\bar{v}_1 \cdot f(\bar{v}_2)$ es igual a $\bar{u}_1 \cdot f(\bar{u}_2)$. Llamando $g \in \mathbb{O}^+(V_2)$ a la rotación tal que $\bar{v}_1 = g(\bar{u}_1)$, se verifica también que $\bar{v}_2 = g(\bar{u}_2)$ pues ambas bases son ortonormales y con la misma orientación. Teniendo además en cuenta que f y g conmutan ($f \circ g = g \circ f$), por ser rotaciones, y que g conserva el producto escalar, resulta que:

$$\bar{v}_1 \cdot f(\bar{v}_2) = g(\bar{u}_1) \cdot f(g(\bar{u}_2)) = g(\bar{u}_1) \cdot g(f(\bar{u}_2)) = \bar{u}_1 \cdot f(\bar{u}_2)$$

3. Si $(\bar{e}_1, \bar{e}_2)$ es una base ortonormal positivamente orientada, como

$$f(\bar{e}_1) = a\bar{e}_1 + b\bar{e}_2 \quad \text{y} \quad f(\bar{e}_2) = -b\bar{e}_1 + a\bar{e}_2$$

se obtiene:

$$\cos \alpha = \bar{e}_1 \cdot f(\bar{e}_1) = \bar{e}_1 \cdot (a\bar{e}_1 + b\bar{e}_2) = a$$
$$\operatorname{sen} \alpha = \bar{e}_1 \cdot f(\bar{e}_2) = \bar{e}_1(-b\bar{e}_1 + a\bar{e}_2) = b$$

EJERCICIO

Demuéstrese que para cualesquiera que sean los ángulos α y β, se verifica que:

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \operatorname{sen} \alpha \operatorname{sen} \beta$$
$$\operatorname{sen}(\alpha + \beta) = \operatorname{sen} \alpha \cos \beta - \cos \alpha \operatorname{sen} \beta$$

RESOLUCION

Las matrices de las rotaciones de ángulos α, β y $\alpha + \beta$ son las:

$$A_\alpha = \begin{bmatrix} \cos \alpha & -\operatorname{sen} \alpha \\ \operatorname{sen} \alpha & \cos \alpha \end{bmatrix}, \quad A_\beta = \begin{bmatrix} \cos \beta & -\operatorname{sen} \beta \\ \operatorname{sen} \beta & \cos \beta \end{bmatrix}$$

$$A_{\alpha+\beta} = \begin{bmatrix} \cos(\alpha + \beta) & -\operatorname{sen}(\alpha + \beta) \\ \operatorname{sen}(\alpha + \beta) & \cos(\alpha + \beta) \end{bmatrix}$$

Como las rotaciones f_α, f_β y $f_{\alpha+\beta}$, de ángulos α, β y $\alpha + \beta$, son tales que $f_{\alpha+\beta} = f_\alpha \circ f_\beta$, resulta que $A_{\alpha+\beta} = A_\alpha A_\beta$. Realizando el producto $A_\alpha A_\beta$ e identificando el resultado con $A_{\alpha+\beta}$ se obtienen las fórmulas del enunciado.

OBSERVACIONES

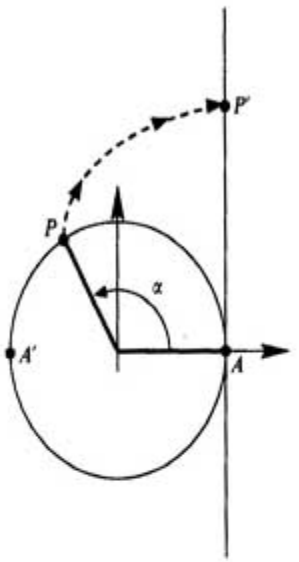


Figura 12.

La idea gráfica e intuitiva que subyace en esto de «medir ángulos» es, poco más o menos, algo del siguiente tipo: imaginemos una circunferencia de radio unidad con centro en el origen O de coordenadas; sea A el punto de la circunferencia situado en el semieje horizontal positivo y sea A' el punto diametralmente opuesto de A. Considérese un ángulo cualquiera $\alpha = \text{áng}\,(\overrightarrow{OA}, \overrightarrow{OP})$, donde P es un punto de la circunferencia.

Supongamos que hubiera un hilo ocupando la posición del arco $\widehat{AP}$, atado en A, y que a dicho hilo se le desenrolla y pasa a situársele hacia arriba, en la vertical que pasa por A; sea P' el punto al que va a parar el extremo del hilo. La distancia $x = AP'$ se va a tomar como medida del ángulo α; llamando $x = \mu(\alpha)$ (medida de α), las cosas deben funcionar de manera que, para ángulos cualesquiera α y α' y si ω y $\bar{\omega}$ son los ángulos nulo y llano:

$$\mu(\alpha + \alpha') = \mu(\alpha) + \mu(\alpha')$$

$$\mu(\omega) = 0 \quad \text{y} \quad \mu(\bar{\omega}) = \pi$$

Supongamos ahora que el hilo sigue uniendo A con P, sobre la circunferencia, pero que lo hace «por abajo», esto es, de manera que ocupe la posición del arco $\widehat{AA'P}$ (el hilo sigue teniendo un extremo atado en A), al desenrollarlo y ponerlo en la vertical (ahora hacia abajo de A). se obtendría otra medida del ángulo α, que valdría $2\pi - x$.

Supongamos ahora que el hilo ocupa la posición inicial del arco $\widehat{AP}$ pero que antes hubiera dado, además, una vuelta completa a la circunferencia; en este caso se obtendría una tercera medida del ángulo α, que valdría $2\pi + x$. Continuando de este modo, obtendremos sucesivas medidas del tipo $x + 2k\pi$, para $k \in \mathbb{Z}$. Por ello, se dice que la medida del ángulo α es el conjunto de todos estos infinitos valores; así, pues, la medida de un ángulo α es un elemento del conjunto cociente $\mathbb{R}/2\pi\mathbb{Z}$, que no es otro que:

$$\mathbb{R}/2\pi\mathbb{Z} = \{x + 2\pi\mathbb{Z}/x \in \mathbb{R}\}, \quad \text{donde} \quad x + 2\pi\mathbb{Z} = \{x + 2k\pi/k \in \mathbb{Z}\}$$

Va a sernos de gran utilidad recurrir a que este conjunto es grupo conmutativo con la siguiente operación (suma):

$$(x + 2\pi\mathbb{Z}) + (y + 2\pi\mathbb{Z}) = (x + y) + 2\pi\mathbb{Z}$$

MEDIDA DE ANGULOS

Las siguientes definiciones vienen a dar respuesta a las consideraciones de tipo intuitivo que acabamos de realizar. Hemos de empezar avisando que, para obtener la función «medida de ángulos», vamos a echar mano de alguna cuestión de análisis matemático, que cae fuera del campo que nos es propio; en concreto, recurriremos a la función «exponencial compleja» y a alguna de sus propiedades:

[416]

Se llama función «medida de ángulos» a un isomorfismo μ, del grupo aditivo de los ángulos $\mathcal{A}$ en el grupo aditivo $\mathbb{R}/2\pi\mathbb{Z}$, tal que $\mu(\bar{\omega}) = \pi$ (*). Para cada ángulo α existe, entonces, un único número real $x \in]-\pi, \pi]$ (que se llama «determinación principal») tal que la medida de α es:

$$\mu(\alpha) = x + 2\pi\mathbb{Z} = \{x + 2k\pi / k \in \mathbb{Z}\}$$

La medida de ángulos es $\mu = \varphi^{-1} \circ \psi \circ \eta \circ \text{rot}$, donde φ, ψ, η y rot son las siguientes aplicaciones, que son todas ellas isomorfismos de los grupos que se indican

$$\mathcal{A} \xrightarrow{\text{rot}} \mathcal{O}^+(V_2) \xrightarrow{\eta} \mathcal{O}^+(2) \xrightarrow{\psi} \mathcal{U} \xrightarrow{\varphi^{-1}} \mathbb{R}/2\pi\mathbb{Z}$$

- $\varphi: \mathbb{R}/2\pi\mathbb{Z} \to \mathcal{U}$, $\varphi(x + 2\pi\mathbb{Z}) = \exp(ix)$, es la exponencial compleja; $\mathcal{U}$ es el grupo multiplicativo de los números complejos de módulo unidad.
- $\psi: \mathcal{O}^+(2) \to \mathcal{U}$ asigna a cada matriz de rotación A (primera columna a, b; segunda columna $-b$, a) el número complejo $\psi(A) = a + bi$, de módulo unidad.
- $\eta: \mathcal{O}^+(V_2) \to \mathcal{O}^+(2)$ asigna a cada rotación su matriz en cualquier base ortonormal positivamente orientada (tal matriz es única).
- rot $\mathcal{A} \to \mathcal{O}^+(V_2)$ asigna a cada ángulo su rotación.

(*) Esta medida se llama «medida en radianes». Si se varía la medida del ángulo llano $\bar{\omega}$, se obtienen otras medidas; al sustituir π por 360 o por 400 se obtienen, respectivamente, las medidas en «grados sexagesimales» y en «grados centesimales».

COMPROBACION

Con excepción de aquello que atañe a la función φ (exponencial), todo cuanto se dice en el anterior enunciado ha sido ya comprobado anteriormente (véase [412] y [413]). Respecto de lo referente a la función φ, lo que se ha hecho es recurrir a resultados conocidos de análisis matemático, que pasamos a comentar:

- La exponencial compleja $x \mapsto \exp(ix)$ es un homomorfismo del grupo aditivo $\mathbb{R}$ en el multiplicativo $\mathcal{U}$, ya que

$$\exp[i(x + x')] = \exp(ix + ix') = [\exp(ix)] \cdot [\exp(ix')]$$

- La aplicación φ está bien definida, esto es, si $x - y \in 2\pi\mathbb{Z}$ (o sea x e y pertenecen a una misma clase de $\mathbb{R}/2\pi\mathbb{Z}$), entonces $\exp(ix)$ es igual a $\exp(iy)$, pues, como $x = y + 2k\pi$ para cierto $k \in \mathbb{Z}$, será:

$$\exp(ix) = \exp(iy + 2k\pi i) = \exp(iy) \cdot \exp(2k\pi i) =$$
$$= \exp(iy) \cdot 1 = \exp(iy), \quad \text{ya que} \quad \exp(2k\pi i) = 1$$

- La aplicación $x \mapsto \exp(ix)$ tiene por imagen (cuando x recorre $\mathbb{R}$) a todo $\mathcal{U}$, con lo que φ es sobreyectiva. Además φ es inyectiva, ya que

$$\varphi(x + 2\pi\mathbb{Z}) = \varphi(y + 2\pi\mathbb{Z}) \Rightarrow \exp(ix) = \exp(iy) \Rightarrow \exp(ix - iy) = 1 \Rightarrow$$
$$\Rightarrow ix - iy = 2k\pi i \quad \text{para cierto } k \in \mathbb{Z} \Rightarrow x + 2\pi\mathbb{Z} = y + 2\pi\mathbb{Z}$$

APENDICE

12 Notas, observaciones y complementos

El carácter especial de este último «capítulo» de apéndices nos ha llevado a incluir, en su final, algunos comentarios y aclaraciones acerca de lo que, en él, ya se ha dicho. En un libro como éste, sobre el «Algebra Lineal», los anteriores Apéndices sólo tiene el objeto de recordar algunos conceptos básicos para que puedan ser utilizados adecuadamente. De ahí que se haya juzgado que lo pertinente era la concisión y, obrando en consecuencia, se ha omitido todo cuanto ha sido posible. Para paliar los abusos cometidos en esta labor de «poda», se incluyen aquí estos complementos:

[417] *NOTA 1: SOBRE ALGUNAS CONJETURAS FAMOSAS*

En matemáticas, se llama conjetura a toda aquella hipótesis, formulada «a priori», acerca de la veracidad o falsedad de cierta propiedad, cuya demostración se ignora. A lo largo de los años, entre la multitud de conjeturas que se han formulado en matemáticas, hay algunas que se han hecho famosas, en las que han trabajado grandes matemáticos sin lograr confirmarlas ni rechazarlas. Citamos aquí algunas de ellas.

a) «Ultimo teorema de Fermat»(*): no existen tres números naturales no nulos x, y, z tales que $x^n + y^n = z^n$ para algún natural $n > 2$. En el año 1637, Fermat afirma tener una demostración del teorema (en una anotación al margen de un libro de Diofanto, dice que ha encontrado una demostración maravillosa del teorema pero que no tiene allí espacio para incluirla), pero no hay pruebas de ello. Los mayores matemáticos de la historia, como Euler, Gauss, Legendre, Dirichlet, Cauchy, Lebesgue, Lamé, Kummer, Kronecker, han trabajado en esta conjetura, obteniendo solamente resultados parciales. En 1976 se había comprobado que era cierta para $n \leqslant 25.000$ y, también, para todo número primo menor que 125.000. En 1984, Faltings prueba que, para cada $n > 2$, no puede haber más que, a lo sumo, un número finito de soluciones. En el año 1908, se ofreció (por F. P. Wolfskehl, en la ciudad alemana de Darmstadt) 100.000 marcos para quién demostrara la veracidad o falsedad de la conjetura; la cantidad ofrecida en este premio está hoy enormemente devaluada. Aunque no tiene excesivo interés la resolución de este famoso «último teorema de Fermat», su im-

(*) El día 23 de junio de 1993, el matemático Andrew Wiles, profesor de la Universidad de Priceton, dio a conocer, en un seminario que se celebraba en Cambridge, una demostración del «último teorema de Fermat». Aunque el resultado está aún por confirmar, todo hace suponer que, al fin, ha quedado probada esta famosa conjetura.

portancia ha sido muy grande en el desarrollo de la aritmética y del álgebra.

b) «Conjetura de Goldbach». En una carta que C. Goldbach envió a Euler en 1742, aquél conjeturaba que todo número natural que fuese par y mayor que 3 era igual a la suma de dos números primos; Euler reconoció que no supo demostrarla. Esta conjetura sigue aún sin resolverse, aunque se han obtenido importantes, resultados parciales (Vinogradov en 1937; Chen en 1973; Vaugham en 1975).

c) «Hipótesis del continuo»: ¿el cardinal de los números reales es el siguiente al de los números enteros o, por el contrario, hay algún otro cardinal entre ellos dos?. Ya Hilbert se hizo esta pregunta, pero el primero en plantearla fue Cantor. Cohen probó en 1963 que ésta es una proposición indecidible, es decir, que es «independiente» de los axiomas de la teoría de conjuntos (véase [419], NOTA 3, *b*).

d) «Problema de lo cuatro colores». Este problema, planteado por primera vez (que se sepa) por un estudiante de Edimburgo hacia el año 1850, pregunta si cuatro colores son suficientes para colorear cualquier mapa geográfico de manera que dos países adyacentes cualesquiera estén pintados de distinto color. Este problema adquirió notoriedad a raíz de que A. Cayley admitiera en 1878 que había intentado resolverlo sin éxito. Desde hace tiempo se sabe que con cinco colores es suficiente para resolver el problema (teorema de Haewood). También se supo pronto que al considerar mapas sobre la cinta de Moebius, sobre la botella de Klein o sobre el plano proyectivo, en lugar de hacerlo sobre el plano o sobre la esfera (que a estos efectos son equivalentes), el problema tenía solución con 6 colores; para el caso del toro se precisa de 7 colores. Durante años, grandes matemáticos han intentado resolver el problema hasta que K. Appel y W. Haken encontraron en 1976 una demostración de que, efectivamente, 4 colores son suficientes.

[418] *NOTA 2: SOBRE LA AUSENCIA DE CONTRADICCIONES EN UNA TEORIA*

a) No es taréa fácil la de comprobar que una teoría matemática no encierra contradicciones, esto es, proposiciones que son verdaderas y falsas a la vez; hoy día no se sabe si la gran mayoría de las teorías matemáticas son o no contradictorias. Hasta el momento no se han encontrado en ellas contradicciones, pero pudiera ocurrir que, cierto día, se encontrara en alguna rama de las matemáticas una proposición que fuese, simultáneamente, cierta y falsa. Si esto llegara a ocurrir, habría que reconsiderar la correspondiente axiomática y, reduciendo o debilitando algunos de los axiomas, buscar el modo de eliminar la contradicción hallada. No es muy probable que, con el tiempo, llegue a demostrarse la coherencia (no contradicción) de, pongamos por caso, la aritmética. No han dado mucho fruto los intentos de demostrar «a priori» que no hay contradicciones en las teorías matemáticas.

b) En el «segundo problema de Hilbert» (segundo de los 23 que en 1900 presentó en el congreso de matemáticas de París) se preguntaba si los

axiomas de la aritmética serían coherentes o si, por el contrario, podrían conducir a alguna contradicción. Gödel probó (en 1931) que no se puede demostrar que los axiomas de la aritmética no conducirán a una contradicción; esto es, que, si la aritmética no es contradictoria, entonces su coherencia (o no contradicción) no es demostrable.

[419] *NOTA 3: SOBRE LAS PROPOSICIONES INDECIDIBLES*

a) El teorema de Gödel asegura que, si una teoría matemática T es no contradictoria e incluye entre sus axiomas a los de la aritmética, entonces no es completa, es decir, en T hay alguna proposición p indecidible. Si a los axiomas de una tal teoría T (por ejemplo, la propia aritmética) se les añade la proposición p (o también la $(no\,(p))$, se obtiene una teoría no contradictoria en la que hay una nueva proposición indecidible. Y así, añadiendo cada vez una nueva proposición, se van obteniendo nuevas teorías, cada vez más restrictivas, que siempre van a tener alguna nueva proposición indecidible.

b) En 1963, Cohen prueba la indecidibilidad de la «hipótesis del continuo» (también llamada conjetura de Cantor o primer problema de Hilbert); esto es, si la teoría de los conjuntos (con el sistema de axiomas de Zermelo-Fraenkel) es no contradictoria, entonces la hipótesis del continuo es indecidible. La hipótesis del continuo dice que no existe ningún conjunto cuyo cardinal esté comprendido entre el cardinal de $\mathbb{N}$ (cardinal de lo numerable) y el cardinal de $\mathbb{R}$ (cardinal del continuo).

c) En 1904, Zermelo formuló el llamado «axioma de elección», que dice: dada una familia de conjuntos no vacíos y disjuntos dos a dos, existe algún conjunto que contiene un único elemento de cada uno de los conjuntos de la familia. Gödel demostró (en 1939) que el «axioma de elección» no es contradictorio con los axiomas de la teoría de conjuntos; Cohen demostró (en 1963) que el «axioma de elección» es indecidible en dicha teoría. Según este axioma, dada una familia de conjuntos no vacíos y disjuntos dos a dos, se puede «elegir» un elemento de cada conjunto de la familia, al que se llama «representante» del conjunto al que pertenece.

[420] *NOTA 4: SOBRE LOS AXIOMAS DE LAS TEORIAS MATEMATICAS*

a) Los axiomas de una teoría son verdades de la misma; son las verdades de las que se deducen las demás verdades de esa teoría. No tiene sentido dudar de los axiomas, del mismo modo que no se duda de las reglas del ajedrez: si se cambian éstas, el juego ya no es el ajedrez, pasa a ser otro juego; así, al alterar los axiomas, se obtiene otra teoría distinta. Los axiomas, que no son sino propiedades ciertas (por definición) de los objetos matemáticos que se contemplan, no han de ser idealizaciones de propiedades de objetos del mundo físico, aunque muchas veces lo sean. Esto último es lo que acontece con los axiomas de la geometría de Euclides; sin embargo, la axiomática de los grupos carece de esta conexión con el mundo sensible.

b) Euclides, al formular en sus «Elementos» (siglo VI a. de C.) los postulados o axiomas de la geometría, incluyó con cierto carácter especial el «Postulado V» (hoy llamado postulado de Euclides o de la paralela) que puede enunciarse diciendo que «por un punto exterior a una recta se puede trazar una y sólo una paralela a dicha recta».

A lo largo de los siglos, muchos matemáticos ilustres intentaron comprobar que no se necesitaba de este postulado por estimar que era consecuencia de los restantes axiomas; estos intentos no dieron fruto. En los primeros años del siglo XIX, Lobachewski y Bolyai, independientemente, construyen una nueva geometría sustituyendo el postulado de Euclides por el postulado «por un punto exterior a una recta se pueden trazar al menos dos rectas paralelas a aquélla (esto es, que no la cortan)». Con ello probó que el postulado de Euclides no puede demostrarse a partir de los otros axiomas, que es una proposición indecidible respecto de éstos; añadiendo a los axiomas de la geometría, en lugar del postulado de Euclides, otro postulado que contradice al de Euclides, se obtiene una geometría, distinta, lógicamente correcta, de la que hoy día hay varios modelos o interpretaciones. Riemann propuso, en 1845, una nueva geometría en la que por un punto exterior a una recta no se puede trazar ninguna paralela a ésta. La geometría de Lobacheski se llama geometría hiperbólica; la de Euclides se llama parabólica; la geometría de Riemann se llama geometría elíptica. Riemann también introdujo una geometría más general, que puede ser elíptica en unos puntos e hiperbólica en otros.

c) Apoyándose en la definición intuitiva de conjunto (agrupación de objetos formando un todo), Cantor hizo sus grandes descubrimientos (a finales del siglo XIX), en el terreno de los conjuntos infinitos, diferenciando las distintas órdenes de infinitud y creando la aritmética de los números transfinitos; es aquí donde Cantor hace su conjetura (hipótesis del continuo). Pero con la teoría intuitiva de conjuntos, al manejar conjuntos infinitos, se encontraron los matemáticos con grandes paradojas como la de Russell, lo que obligó a axiomatizar la teoría de conjuntos; fundamentalmente, hay dos sistemas de axiomas, para la teoría de conjuntos, uno debido a Zermelo y Fraenkel y otro de Neumann, Bernayo y Gödel (este último sistema distingue entre clases y conjuntos, donde éstos son unas clases particulares: aquéllas que pueden ser elementos de otras clases). Con estas axiomáticas se obtiene la «teoría restringida de conjuntos»; si a los anteriores axiomas se les añade el axioma de elección (véase [419], NOTA 3, *c*), se obtiene la «teoría estándar de conjuntos».

d) Comentaremos aquí, someramente, uno de los sistemas de axiomas de la teoría de conjuntos; nos vamos a referir al sistema de Zermelo-Fraenkel (fue Zermelo el primero que lo formuló, en 1905; Fraenkel lo modificó en 1920). Este consta de 9 axiomas, que se refieren a:

- El 1.º (llamado de extensionalidad) viene a decir que dos conjuntos son iguales si y sólo si tienen los mismos elementos.
- El 2.º (llamado del conjunto vacío) viene a decir que existe un conjunto que carece de elementos (el conjunto vacío, $\varnothing$).
- El 3.º (llamado de emparejamiento) viene a decir que dados dos conjuntos A y B, existe un conjunto cuyos elementos son A y B, al que se denota por $\{A, B\}$.

- El 4.° (llamado de la unión) viene a decir que, si A es un conjunto de conjuntos, existe un conjunto que tiene por elementos a los elementos de los conjuntos de A.
- El 5.° (llamado del infinito) asegura la existencia de un conjunto infinito. Distinguir entre A y $\{A\}$, permite construir los conjuntos $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, ...
- El 6.° (llamado axioma de reemplazamiento) viene a significar que, para cualquiera que sea la proposición p, existe el conjunto de los elementos que satisfacen a p.
- El 7.° (llamado del conjunto potencia) viene a decir que para cada conjunto A existe el conjunto B de las partes de A. Como B resultará tener mayor cardinal que A, se pueden obtener conjuntos cuyo cardinal es cada vez mayor.
- El 8.° es el axioma de elección, del que ya hablamos en la nota 3, *c*.
- El 9.° (llamado de regularidad) impide que se presenten casos indeseables como que un conjunto sea un elemento de sí mismo.

[421] *NOTA 5: SOBRE LAS PARADOJAS DE CANTOR Y RUSSELL*

Según se señaló en [420], NOTA 4, *c*, la teoría intuitiva de conjuntos conduce a paradojas, cosa no deseable que se evita con la debida axiomatización de esta teoría. De entre estas paradojas, destacan las de Cantor y Russell, de las que vamos a hablar aquí.

a) Antes de referirnos a la paradoja de Cantor, comencemos recordando el teorema de Cantor: para cualquiera que sea el conjunto C, el cardinal de C es menor que el cardinal del conjunto $\mathcal{P}(C)$, de las partes de C.
La paradoja de Cantor arranca de suponer que existe el conjunto C de todos los conjuntos. De ser así, todos los subconjuntos de C serían, como cualquier conjunto, elementos de C, luego $\mathcal{P}(C)$ habría de ser un subconjunto de C y, consecuentemente, el cardinal de $\mathcal{P}(C)$ sería menor o igual que el cardinal de C, lo que contradice al teorema de Cantor.

b) Podría pensarse que la contradicción que acabamos de detectar, en la paradoja de Cantor, podría estar motivada por haber admitido que un conjunto puede ser, al tiempo, uno de los elementos que lo forman. Ello nos llevaría a considerar solamente conjuntos que no sean elementos de sí mismos, que se llaman normales, rechazando los anormales, que son los que sí son elementos de sí mismos. Esto da pie a la paradoja de Russell: Sea R el conjunto de todos los conjuntos normales. ¿Es R un conjunto normal? Si R fuese un conjunto normal, entonces por la propia definición de R resultaría que R habría de ser uno de los elementos de R, lo que significa que R tendría que ser un conjunto anormal, contra lo supuesto. Si, por el contrario, R fuese un conjunto anormal, entonces por la propia definición de R resultaría que R no podría ser un elemento de R, lo que significa que R habría de ser un conjunto normal, contra lo supuesto. Nos encontramos con que R no puede ser, pues, ni normal ni anormal, que es una contradicción manifiesta.
Esta paradoja es, en esencia, análoga a las llamadas paradojas del barbero y de los catálogos:

- Paradoja del barbero: en un pueblo, el barbero afeita a todos los hombres que no se afeitan a sí mismo, ¿quién afeita al barbero?
- Paradoja de los catálogos: un bibliotecario desea hacer el catálogo C de los catálogos que no están catalogados. ¿El catálogo C deberá estar entre los catálogos que se catalogan en este nuevo catálogo?

[422] *NOTA 6: SOBRE LOS OPERADORES LOGICOS*

Para «componer» dos proposiciones, que llamaremos p y q, se dispone de las conectivas (u operadores lógicos), que representaremos genéricamente por $*$; las proposición compuesta se denotará poniendo $p*q$ De entre las conectivas, solamente se han considerado en el texto el «*o*» (disyunción inclusiva), el «*y*» (conjunción), el $\rightarrow$ (condicional) y el $\leftrightarrow$ (bicondicional), pero en total hay 16 conectivas. Hablemos algo de ellas.

Una conectiva $*$ queda definida por el valor de verdad que le corresponde a $p*q$, en función de los valores de verdad de p y de q; esto es, conocer $*$ es tanto como conocer el valor de verdad que se asigna a $p*q$ según que los de p y q sean v y v, sean v y f, sean f y v o sean f y f. Como para cada uno de estos cuatro supuestos, a $p*q$ se le puede asignar el valor v o el valor f, resulta que hay un total de $2^4 = 16$ conectivas. Además de las cuatro conectivas que ya se han considerado, cabe citar a las siguientes:

- La *disyunción excluyente* de las proposiciones p y q, que representaremos poniendo $p\ o^*\ q$, es la proposición que es verdadera cuando lo es una y sólo una de las p y q. Nótese que

$$p\ o^*\ q = no(p \leftrightarrow q)$$

  ya que ambas proposiciones tienen los mismos valores de verdad.
- La *incompatibilidad* de las proposiciones p y q, que representaremos poniendo $p|q$, es la proposición que es falsa si p y q son ambas verdaderas y que es verdadera en los demás casos. La incompatibilidad, al igual que la disyunción excluyente, puede expresarse en función de las anteriores conectivas, ya que

$$p|q = no(p\ y\ q)$$

  como se comprueba viendo que los valores de verdad de estas dos proposiciones son los mismos.

[423] *NOTA 7: SOBRE LOS TEOREMAS Y LAS IMPLICACIONES*

En una teoría matemática, cuyos axiomas representaremos por a, para establecer la veracidad de una proposición p, se recurre a los axiomas a y a teoremas ya conocidos, que representaremos por t. La proposición p será verdadera si y sólo sí lo es la condicional $(a\ y\ t) \rightarrow p$; esto es, p será un teorema si y sólo sí $(a\ y\ t) \Rightarrow p$. Es por ello que los términos «teorema» e «implicación» se suelen

considerar equivalentes (téngase en cuenta que las implicaciones, por ser proposiciones verdaderas, son teoremas).

Los teoremas se suelen enunciar en forma de implicaciones. Así, por ejemplo, el teorema «para todo número natural n, se verifica que $2^n \geq n+1$» puede enunciarse en forma de implicación diciendo «si n es un número natural, entonces $2^n \geq n+1$». El teorema «hay infinitos números primos» puede expresarse como implicación en la forma «si P es el conjunto de los números primos, entonces P es infinito».

Cuando, en el contexto de una teoría matemática, se enuncia un teorema en la forma $p \Rightarrow q$, se da por supuesto que también forman parte de la hipótesis todos los axiomas de la teoría, que representaremos por a, y los teoremas ya demostrados, que representaremos por t; por ello. en puridad, el teorema debiera haberse formulado poniendo $(a\ y\ t\ y\ p) \Rightarrow q$.

[424] *NOTA 8: SOBRE LA «INDUCCION», EL AXIOMA DE ZORN Y EL TEOREMA DE ZERMELO*

a) El llamado principio de *inducción completa* establece que, si una cierta propiedad, que depende de la variable natural n, se verifica para cierto $n_0 \in \mathbb{N}$ y, en el supuesto de que se verifica para los valores n_0, n_0+1, n_0+2, ..., $n-1$, también se verifica para el valor n, entonces la propiedad se verifica para todos los números naturales mayores que n_0. Esto es, dada una función proposicional p, con variable x que recorre el conjunto

$$A = \{n \in \mathbb{N} / n \geq n_0\}$$

se verifica que:

$$\left.\begin{array}{l} p(n_0) \\ [\forall\, h \in \{n_0,\ n_0+1,\ \ldots,\ n-1\},\ p(h)] \ \Rightarrow\ p(n) \end{array}\right\} \ \Rightarrow\ \forall\, n \in A,\ p(n)$$

El principio de inducción completa es un caso particular del principio de *inducción trasfinita*, en el que $\mathbb{N}$ se sustituye por un *conjunto bien ordenado* (conjunto ordenado en el que todo subconjunto no vacío tiene mínimo). Este principio (o teorema, dependiendo del punto de partida) dice que: si $(A, \leqslant)$ es un conjunto bien ordenado y si p es una función proposicional, con una variable x que recorre A, se verifica que:

$$[x \in A,\ (\forall\, y \in \{a \in A / a < x\},\ p(y)) \ \Rightarrow\ (p(x))] \ \Rightarrow\ [\forall\, x \in A,\ p(x)]$$

b) Para poder aplicar el principio de inducción transfinita se precisa, pues, de una buena ordenación. La existencia de una buena ordenación queda garantizada por el *Teorema de Zermelo* que, apoyándose en el «axioma de elección» (véase [419], NOTA 3, *c*), asegura que «todo conjunto admite una buena ordenación». Conviene observar que este teorema no proporciona la buena ordenación cuya existencia asegura. No siempre se sabe cómo «ordenar bien» un conjunto dado.

c) El «axioma de elección» (véase [419], NOTA 3, *c*) resulta ser equivalente al *axioma de Zorn* (que se enuncia a continuación) y también al teorema

de Zermelo. Admitido, pues, uno de estos tres (como axioma), se puede demostrar la veracidad de los otros dos, que pasan a ser teoremas. El axioma de Zorn (o teorema de Zorn, según se mire) dice que: «en un conjunto ordenado $(A, \leqslant)$, si toda cadena (esto es, subconjunto totalmente ordenado) de A está acotada superiormente, entonces A tiene un elemento maximal, al menos.

[425] *NOTA 9 SOBRE LOS CONJUNTOS FINITOS E INFINITOS*

a) Hay bastantes propiedades relativas a los conjuntos finitos que, no obstante su apariencia inocente, se demuestran con mayor dificultad de la que cabría esperar. Esto es lo que ocurre con la siguiente afirmación: «si A es un conjunto finito y B es un conjunto estrictamente incluido en A, entonces no existe ninguna biyección entre A y B». Comprobémoslo:

- Sean $n = \text{card } A$ y $N = \{1, 2, ..., n\}$; sabemos que hay una biyección $\varphi: A \to N$ y que, por tanto, B y $M = \varphi(B)$ son equipotentes. Nótese que M está estrictamente incluido en N. Lo que se pretende comprobar equivale a esto otro: no existe una biyección entre N y M.
- Si fuese $n \in M$ (que puede no serlo), como existe algún $k \in N$ tal que $k \notin M$, entre los conjuntos M y $M' = M \cup \{k\} - \{n\}$ podemos establecer la biyección $\psi: M \to M'$, $\psi(i) = i$ para $i \neq n$ y $\psi(n) = k$. Por tanto, entre N y M existirá o no una biyección si, y sólo si, entre N y M' existe o no, respectivamente, una biyección; consecuentemente, se puede suponer que $n \notin M$, cosa que se hará a partir de aquí.
- La propiedad a demostrar es evidente para $n = 1$; procediendo por inducción, se supondrá que la propiedad es cierta para Card $A = 1, 2, ..., n - 1$ y habrá que comprobar que también lo es para Card $A = n$. Supongamos que esto último no es cierto, es decir, que existe una biyección $f: N \to M$ para cierto $M \subset N$ con $n \notin M$. Llamando entonces $h = f(n)$, $M_1 = M - \{h\}$ y $N_1 = N - \{n\} = \{1, 2, ..., n - 1\}$, la aplicación $g: N_1 - M_1$ con $g(i) = f(i)$ para todo $i \in N_1$ será también una biyección; ahora bien, como M_1 está estrictamente incluido en N_1, este último resultado contradice a la hipótesis de inducción, luego la propiedad es cierta.

b) Comprobemos ahora que los conjuntos no finitos son aquellos y sólo aquellos que se pueden poner en correspondencia biyectiva con alguno de sus subconjuntos propios. Más exactamente, que son equivalentes las tres afirmaciones:

1.ª El conjunto A es infinito (o sea, no es finito).
2.ª El conjunto A incluye a un conjunto infinito numerable.
3.ª El conjunto A es equipotente a alguna de sus partes propias.

Todo ello lo comprobaremos viendo que 1.ª $\Rightarrow$ 2.ª, que 2.ª $\Rightarrow$ 3.ª y que 3.ª $\Rightarrow$ 1.ª. Así es, en efecto:

(1.ª $\Rightarrow$ 2.ª) Como es $A \neq \emptyset$, existe $a_0 \in A$; como por ser A infinito es $A - \{a_0\} \neq \emptyset$, existe $a_1 \in A$ con $a_1 \neq a_0$; en el supuesto de que existen a_0,

$a_1, ..., a_{n-1} \in A$ con $a_i \neq a_j$ para $i \neq j$, por ser A infinito es $A - \{a_0, a_1, ..., a_{n-1}\} \neq \emptyset$ y, por tanto, existe $a_n \in A$ con $a_n \neq a_i$ para $i = 0, 1, ..., n-1$. El principio de inducción completa permite, pues, afirmar que para cada $n \in \mathbb{N}$ existe un $a_n \in A$ tal que $a_n \neq a_m$ para todo $m \neq n$. Llamando $A' = \{a_0, a_1, ..., a_{n-1}, ...\}$, la aplicación $f: \mathbb{N} \to A'$, $f(n) = a_n$, es una biyección, luego A' es infinito numerable, con lo que concluye la demostración, ya que $A' \subset A$.

($2.^a \Rightarrow 3.^a$) Se sabe que, para cierto conjunto $A' \subset A$, existe una biyección entre $\mathbb{N}$ y A', es decir, que A' es del tipo $A' = \{a_n / n \in \mathbb{N}\}$ con $a_i \neq a_j$ para $i \neq j$. Considérese la aplicación $f: A \to A - \{a_0\}$ dada por: si $a \in A - A'$ es $f(a) = a$; $f(a_n) = a_{n+1}$ para $n \in \mathbb{N}$. Como f es evidentemente biyectiva, resulta que A es equipotente a $f(A) = A - \{a_0\}$, que es una parte propia de A.

($3.^a \Rightarrow 1.^a$) Como un conjunto finito no es equipotente a ninguna de sus partes propias y A lo es a una de ellas, resulta que A no es finito, luego es infinito.

c) *Teorema de Berstein:* dados dos conjuntos A y B, si existen dos aplicaciones inyectivas $f: A \to B$ y $g: B \to A$, entonces A y B son equipotentes. Para probar este teorema, puede seguirse el siguiente camino. Para cada $a \in A$, la imagen recíproca $g^{-1}(a)$ o es vacía o consta de un solo elemento de B; en este segundo supuesto, $f^{-1}(g^{-1}(a))$ o es vacío o consta de un solo elemento de A; etc. Sean A', A_a y A_b los subconjuntos de A formados por aquellos elementos para los que el anterior proceso no concluye nunca, concluye en un elemento de A y concluye en un elemento de B, respectivamente. De igual modo se definen B', B_a y B_b. Se verifica que $f(A') = B'$ (y que $g(B') = A'$), que $f(A_a) = B_a$ y $g(B_b) = A_b$, con lo que $A' \simeq B'$, $A_a \simeq B_a$ y $A_b \simeq B_b$; como $\{A', A_a, A_b\}$ y $\{B', B_a, B_b\}$ son particiones de A y B, respectivamente, se concluye que $A \simeq B$.

d) *Teorema de Cantor:* para cualquiera que sea el conjunto A, se verifica que Card $A <$ Card $\mathcal{P}(A)$. Este teorema se comprueba sin grandes dificultades siguiendo los siguientes pasos (reducción al absurdo): se supone que existe una aplicación sobreyectiva $f: A \to \mathcal{P}(A)$; se considera el conjunto $B = \{x \in A / x \notin f(x)\}$; por ser f sobreyectiva y $B \in \mathcal{P}(A)$, existe $b \in A$ tal que $f(b) = B$; este $b \in A$ o es de B o no es de B; si $b \in B$, por la definición de B, habrá de ser $b \notin f(b) = B$, que es una contradicción; si $b \notin B$, por la definición de B, habrá de ser $b \in f(b) = B$, que es una contradicción; si $b \notin B$, por la definición de B, habrá de ser $b \in f(b) = B$, que es una contradicción. Por ello, no puede haber una biyección de A en $\mathcal{P}(A)$.

Alfabeto griego

Figura		Nombre	Figura		Nombre
A	α	alfa	N	ν	nu, ny
B	β	beta	Ξ	ξ	xi
Γ	γ	gamma	O	o	ómicron
Δ	δ	delta	Π	π	pi
E	ε	épsilon	P	ρ	rho
Z	ζ	dseta	Σ	σ, ς	sigma
H	η	eta	T	τ	tau
Θ	ϑ, θ	theta	Y	υ	ýpsilon
I	ι	iota	Φ	φ	fi
K	κ	kappa	X	χ	ji
Λ	λ	lambda	Ψ	ψ	psi
M	μ	mu, my	Ω	ω	omega

Indice

Ampliación de una aplicación, 712
Angulo,
  de dos planos, 471
  de dos rectas, 443, 470
  de dos vectores, 264, 774
  de recta y plano, 471
  de semirrectas, 772
  llano, 774
  nulo, 774
Angulos rectos, 774
Anillo,
  conmutativo, 750
  de endomorfismos, 182
  de integridad, 760
  de matrices, 183
  unitario, 756
Anillos, 756
Aplicación,
  bilineal, 210
  canónica, 720
  constante, 711
  creciente, 725
  dual o traspuesta, 193
  entre conjuntos, 711
  idéntica, 711
  identidad, 151
  inclusión, 711
  inyectiva, 713
  lineal, 150
    asociada a una forma cuadrática, 220
    inyectiva, 156
  nula, 151
  ortogonal, 288
  proyección, 711
  recíproca de otra, 713
  sobreyectiva, 713
Aplicaciones,
  monótonas, 725
  parciales, 220
Area de un triángulo, 448, 480
Asíntotas de la hipérbola, 557, 622
Automorfismo, 742, 748, 761
  en un espacio vectorial, 158
  ortogonal, 290
Automorfismos de un espacio vectorial, 189
Autovalores,
  de un endomorfismo, 327
  de una matriz, 330
  de una transformación ortogonal, 364
Autovector de un endomorfismo, 327
Autovectores de una matriz, 330
Axioma,
  de Zorn, 784
  de elección, 780, 784, 791
Axiomas, 691, 780
  de Zermelo-Fraenkel, 780, 781

Baricentro, 522, 527
Base,
  cambio de, 135
  definición de, 127
  dual, 192
  ortogonal, 270
Bases, existencia de, 129
Bicondicional, 695
Bisectrices, 544
Bloques,
  de Jordan, 385, 399
  de una matriz, 60
Buena ordenación, 784

Cajas de Jordan, 385, 399
Campo de definición, 711
Característica de un cuerpo, 767
Cardinal de un conjunto, 728
Centro,
  de un elipsoide, 575
  de una cónica, 619
  de una cuádrica, 662
Centros de los hiperboloides, 578, 581
Cilindro,
  cuadrático, 693
  elíptico, 660
  hiperbólico, 660
  imaginario, 660
  parabólico, 660
Circunferencia, 449

Clases de equivalencia, 719
Clasificación, afín, 614, 656
de las cónicas, 614
de las cuádricas, 656
general,
de las cónicas, 617, 618
de las cuádricas, 660
Coeficiente de Fourier, 283
Combinaciones lineales de vectores, 26, 118, 120
Complesificación de un espacio vectorial real, 342
Composición,
de aplicaciones, 714
de homomorfismo, 179
Condición,
necesaria, 700
suficiente, 700
Condicional, 695
Conectivas, 695
Congruencia,
módulo p, 719, 739, 760
ortogonal, 361
Cónica de tipo,
elíptico, 614
hiperbólico, 614
parabólico, 614
Cónica del infinito de una cuádrica, 656
Cónicas, 600
con centro, 619
degeneradas, 608
imaginarias, 611
ordinarias, 608
reales, 611
Conjetura de Goldbach, 779
Conjeturas matemáticas, 778
Conjunción de proposiciones, 695
Conjunto,
acotado, 725
bien ordenado, 784
cociente, 720
complementario de otro, 705
de las partes de otro, 704
de llegada, 711
de verdad, 704
infinito numerable, 730
intuitivo, 702
ordenado, 724
origen de una aplicación, 711
universal, 703
vacío ($\varnothing$), 703
Conjuntos,
coordinables, 728
disconjuntos, 706
equipotentes, 728
finitos, 729, 785
infinitos, 730, 785
Cono,
asintótico de un hiperboloide, 579, 582, 667
cuadrático, 693
imaginario, 660
real, 660
tangente a una cuádrica, 589, 647
Contradicciones, ausencia de, 779
Contraejemplo, 701
Coordenadas,
baricéntricas, 525
cambio de, 135
cartesianas, 429, 505
absolutas, 486
cambio de, 430, 507
contravariantes, 265
covariantes, 265
de vectores, 132
homogéneas, 486
cambio de, 490
de los puntos del infinito, 493
Correspondencia, 710
Coseno de un ángulo, 269, 774
Cosenos directores, 264
Cotas superior e inferior, 725, 726
Criterio de Sylvester, 251
Cuádricas, 632
con centro, 662
degeneradas, 693
de tipo, 657
elíptico, 657
hiperbólico, 657
parabólico, 657
imaginarias, 646
no regladas, 597, 653
ordinarias, 693
reales, 646
regladas, 594, 653
Cuadrilátero completo, 540
Cuantificador,
existencias, 698
universal, 698
Cuerpo algebraicamente cerrado, 340
Cuerpos, 765

Deducción lógica, 691
Demostración, 701
por reducción al absurdo, 701
Dependencia lineal, 26, 122
de puntos, 523
Descomposición canónica de una aplicación, 722
Desigualdad,
de Bessel, 284
de Minkowski, 259

Desigualdad *(Cont.)*
de Schwarz, 259
triangular, 428, 503
Determinante,
de la matriz inversa, 86
de un endomorfismo, 174
de una matriz,
cálculo del, 83
definición, 72
expresión, 77
propiedades, 73, 74
de Vandermonde, 81
del producto de dos matrices, 85
desarrollo para una línea, 88
Diagonalización,
de una forma cuadrática, 232, 236
en bloques, 369, 378, 383
de un endomorfismo, 378
de una transformación ortogonal, 369
triangulares, 383
ortogonal, 353
de un endomorfismo simétrico, 359
de una forma cuadrática, 362
de una matriz simétrica, 359
por congruencia de una matriz, 232, 235
por equivalencia, 169
Diagramas de Venn, 706
Diámetro conjugado, 565
Diámetros de las cónicas, 564
Diferencia de conjuntos, 705
Dimensión de un espacio afín, 505
Dirección,
de un paraboloide, 665
de un plano, 432
de una parábola, 621
de una recta, 431
Directriz,
de la parábola, 560
de una cónica, 568
Distancia, 428, 503, 504
de un punto a plano, 472
de punto a recta, 445, 434
de punto a variedad lineal, 529
entre dos rectas, 476
entre variedades lineales, 534
Divisores de cero, 52, 67, 759
Disyunción,
de proposiciones, 659
excluyentes, 783

Ecuación,
característica, 332, 334
de un hiperplano, 515
de una cónica, 600
de una cuádrica, 632
focal de una cónica, 570
homogénea, 605, 638
de una cónica, 605
de una cuádrica, 638
normal, 447, 479
de un plano, 479
de una recta, 447
reducida, 553, 555, 560, 573, 627, 629, 673
de una elipse, 553
de la hipérbola, 555
de la parábola, 560, 629
de un elipsoide, 573
de un paraboloide, 674
de una cónica con centro, 627
de una cuádrica con centro, 673
trinomía de una cónica, 561
Ecuaciones,
de las rectas, 434, 441, 456, 458
de los planos, 453, 455
de rectas y planos ampliados, 495
de una aplicación lineal, 162
de una variedad lineal, 517
homogéneas de rectas y planos, 487, 488
reducidas, 577, 580, 583, 585, 674
de los hiperboloides, 577, 580
de los paraboloides, 583, 585, 674
Eje,
de la parábola, 560, 629
de un paraboloide, 583, 585, 674
Ejes,
de la elipse, 553
de la hipérbola, 555
de la cónica con centro, 627
de las cuádricas con centro, 673
de los hiperboloides, 578, 581
de un elipsoide, 573, 575
Elemento,
fijo, de una aplicación, 712
idempotente, 737
inverso de otro, 737, 758, 765
neutro, 736, 745, 756
nulo, 737
opuesto de otro, 737, 756
simétrico de otro, 736, 745
unidad, 737, 756
Elementos,
invertibles de un anillo, 758
regulares, 736
Eliminación de incógnitas, 16
Elipse, 553, 615
de garganta, 578
imaginaria, 617, 618
real, 617, 618

Elipsoide, 573, 657
imaginario, 660
real, 660
Endomorfismo, 742, 748, 761
diagonalizable, 346, 347
simétricos, 353
triangulares, 379
Enunciados, 694
Epimorfismo, 748
Equivalencia asociada a una aplicación, 722
Escalares, 23, 116
Esfera,
de E_n, 533
de E_3, 482
de Monge, 678
Espacio,
afín, 427, 499, 500
ampliado, 492
euclídeo, 428, 502
bidual, 195
dual, 190
métrico, 504
proyectivizado, 492
proyectivo, 497
vectorial, 116
cociente, 186
euclídeo, 254
euclídeo usual, 254
normado, 262
Espacios,
puntuales, 499
vectoriales isomorfos, 158
Estructura algebraica, 734
Excentricidad de una cónica, 568
Expresión,
canónica de una forma cuadrática, 246
del producto escalar, 256
Expresiones lógicas, 697
Extremos,
de un conjunto, 726
de una función, 727

Foco
de la parábola, 560
de una cónica, 568
Focos,
de la elipse, 553
de la hipérbola, 555
Forma,
bilineal, 209
antisimétrica, 215
simétrica, 215
canónica, 369, 373, 391, 395
de Jordan, 391, 395
de una transformación ortogonal, 369, 373
cuádrica, 218
degenerada, 229
ordinaria, 229
diagonal, 350
de un endomorfismo, 350
de una matriz, 350
Formas cuadráticas,
definidas, 240, 248
semidefinidas, 240, 248
Fórmulas de las dimensiones, 147, 521
Función,
acotada, 727
de un conjunto en otro, 711
proposicional, 598

Generatrices rectilíneas, 594, 596, 653
Geometría,
de Lobacheski o hiperbólica, 781
de Riemann o elíptica, 781
euclídea o parabólica, 781
Grados de libertad, 20
Gráfica o grafo, 710, 716
Grupo, 745
abeliano, 745
de sustituciones, 752
lineal, 184
multiplicativo de un anillo, 758
ortogonal, 287
simétrico, 752

Haces,
de planos, 464
de rectas, 440
Hipérbola, 555, 615
conjugada de otra, 557
equilateral, 555
Hiperboloide de una hoja,
elíptico o no reglado, 580, 660
reglado o hiperbólico, 577, 660
Hiperboloides, 657
Hipercono tangente a una esfera, 537
Hiperplano, 509, 515
radical, 545
tagente a una esfera, 534
Hipótesis, 700
del continuo, 733, 779
Homomorfismo,
de anillos, 761
de cuerpos, 768
de espacios vectoriales, 150
de grupos, 748
entre estructuras algebraicas, 740
inyectivo, 751

Ideales de un anillo, 763
Imagen,
de un homomorfismo, 750, 762
de una aplicación, 711
lineal, 153
homomorfa, 742, 762
recíproca, 711
recíproca de un vector, 187
Implicación, 700, 783
Incompatibilidad de proposiciones, 768
Independencia lineal, 29, 122
de puntos, 523
Indices,
libres, 5
mudos, 5
Inclusión, 704
Inducción,
completa, 784
transfinita, 784
Infimo de un conjunto, 726
Intersección,
de cónica y recta, 562, 606
de conjuntos, 706
de cuádrica y plano, 641
de cuádrica y recta, 586, 640
de dos esferas, 535
de plano, 458, 462
de subconjuntos, 748
de subespacios vectoriales, 119
de variedades lineales, 518
Intersecciones de rectas, 438, 468, 469
y planos, 466, 467
Invariante,
afín de una cónica, 603
afín de una cuádrica, 636
proyectivo,
de una cónica, 603
de una cuádrica, 636
Invariantes métricos,
de una cónica, 625
de una cuádrica, 671
Isomorfismo,
de coordenadas, 133
de grupos, 748
entre espacios vectoriales, 158
entre estructuras algebraicas, 742
respecto del orden, 725

Ley,
de composición, 734
de inercia, 244
de Morgan, 707, 708

Matriz,
antisimétrica, 43
canónica, 41, 169, 247
de congruencia, 247
de equivalencia, 41, 169
cuadrada, 43
de Gram, 211, 256
de Jordan, 385
de los términos cuadráticos, 600, 632
de paso, forma canónica de Jordan, 407
de un sistema de ecuaciones, 6
de una aplicación lineal, 163, 167
de una forma,
bilineal, 211
cuadrática, 221
diagonal, 43
diagonalizable, por semejanza, 346, 347
escalonada, 13
inversa,
expresión, 91
de otra, 62
invertible, 62, 67
métrica, 256
nula, 45
opuesta, 45
ortogonal, 290
regular, 67
simétrica, 43
singular, 67
traspuesta de otra, 54
triangular, 43
unidad, 52
Matrices,
congruentes, 213, 221, 247
de endomorfismos, 173
definidas, 243, 249
de las cónicas, 600, 601, 623
de las cuádricas, 632, 634, 635
equivalentes, 41, 167, 170
ortogonales, 295
directas, 295
inversas, 295
semejantes, 173
semidefinidas, 243, 249
triangularizables, 379
Máximo de un conjunto, 726
Medida de ángulos, 777
Menores de una matriz, 93
Método de eliminación,
de Gauss, 18
de Gauss-Jordan, 18
Mínima distancia, 281
Mínimo de un conjunto, 726
Modelos matemáticos, 693
Monomorfismo, 748

Multiplicación de matrices por bloques, 61
Multiplicidad,
algebraica, 336
en $\mathbb{R}$ y $\mathbb{C}$, 341
geométrica, 336

Negación de una proposición, 695
Norma de un vector, 259
Núcleo,
de un homomorfismo, 750, 762
de una aplicación lineal, 152
de una forma cuadrática, 229
Números,
enteros, 722
naturales, 729

Objetos matemáticos, 691
Operación,
asociativa, 736
cociente, 739
compatible con una equivalencia, 739
conmutativa, 736
distributiva respecto de otra, 736, 756
externa, 735
interna, 734
Operaciones,
booleanas, 707, 708
elementales, 9, 12, 33
en un sistema de ecuaciones, 9
en un sistema de vectores, 33
en una matriz, 12
Operadores lógicos, 783
Orden,
parcial, 724
total, 724
Orientación de una base, 174
Ortonormalización de Gram-Schmidt, 273

Parábola, 560, 615
Paraboloide,
elíptico o no reglado, 583, 660
hiperbólico o reglado, 585, 660
Paraboloides, 657
Paradoja,
de Cantor, 782
de los catálogos, 783
de Russell, 782
del barbero, 783
Parámetro de una parábola, 560
Parámetros de una cónica, 571
Paralelismo,
de planos, 461
de recta y plano, 461
de rectas, 436, 461, 468, 469
de variedades, 511
Par de objetos matemáticos, 709
Parte estable,
para una aplicación, 712
para una operación, 712
Partición de un conjunto, 706
Permutaciones, 76
Perpendicular,
a una variedad, 528
común a dos rectas, 478
Perpendicularidad,
de planos, *460*
de rectas, 444, 460
de rectas y planos, 460
Pivote, 18
Plano,
afín, 427
afín-euclídeo, 428
del infinito, de $\bar{E}_3$, 496
polar, 589, 647
tangente a una cuádrica, 591, 641, 647
tangente a una esfera, 482, 483
Planos, 432, 452, 510
principales, 575, 578, 581, 583, 585
de los hiperboloides, 578, 581
de los paraboloides, 583, 585
de un elipsoide, 575
Polar respecto a una cónica, 567, 612
Polinomio característico,
de un endomorfismo, 334
de una matriz, 332
Polo de un plano, 680
Postulado V° de Euclides, 781
Potencia,
del continuo, 730, 732
de lo numerable, 730, 732
respecto de una esfera, 545
Problema de los cuatro colores, 779
Producto,
cartesiano, 709
de matriz por escalar, 45, 177
de matrices, 49, 179
de vector por escalar, 23, 116
escalar, 254
canónico, 254
mixto, 302
vectorial, 308, 311
Proposiciones, 694
indecibles, 780
Proyección, 710
ortogonal, 281
sobre un plano, 472
sobre una recta, 445, 474
sobre una variedad, 528

Punto,
medio, 431, 527
unidad, 493
Puntos, 427
conjugados respecto de una cuádrica, 590
del infinito, 491
de una cónica, 641
de una cuádrica, 656
dependientes, 523
elípticos en las cuádricas, 650
fundamentales, 493
impropios, 491
independientes, 523
parabólicos en conos y cilindros, 650
singulares de cónicas, 608
de cuádricas, 643

Radiación de planos, 465
Raíces de polinomios, 339
Rango,
de menores, 94
de un sistema de vectores, 30, 137
de una aplicación lineal, 153
de una forma cuadrática, 221
de una matriz, 37
Razón simple, 544
Recta del infinito, de $\bar{E}_2$, 496
Rectángulo principal de una hipérbola, 557
Rectas, 430. 452, 510
perpendiculares, 44
Recubrimiento. 706
Referencia.
afín, 525
baricéntrica. 525
cartesiana. 429. 505
rectangular. 429. 509
Refutar, 701
Regla de Cramer. 98
Relación,
binaria. 716
antisimétrica. 717
reflexiva. 717
simétrica. 717
transitiva. 717
de conjugación. 225
de Chasles. 427
para ángulos, 772
de equivalencia. 719
de orden. 724
Restricción,
de una aplicación. 712
Rotaciones, 293
bidimensionales. 297
tridimensionales. 300

Secciones,
cíclicas, 576
cónicas, 571
Segmento, 431
Segundo problema de Hilbert, 779
Semejanza ortogonal, 361
Seno de un ángulo. 774
Seudoejes de un paraboloide, 674
Signatura,
de una forma cuadrática, 244
de una matriz simétrica, 245
Signo,
pertenencia ($\in$), 702
de una sustitución, 76
sumatorio, 5
Símbolos de Kronecker, 90
Simetría ortogonal. 293
bidimensional, 297
Simetrías ortogonales tridimensionales, 300
Sistema
de ecuaciones lineales, 4
escalonado, 13, 14
de vectores linealmente dependientes, 26, 122
generador, 120
ortogonal de vectores, 267
Sistemas,
de Cramer, 98
de ecuaciones equivalentes, 9
de ecuaciones lineales homogéneas, 7
de vectores linealmente independientes, 29, 122
equivalentes de vectores, 121
Subanillos. 758
Subconjunto. 704
Subcuerpo. 766
Subespacio,
conjugado en un vector. 226
ortogonal a otro. 278
vectorial. 118
Subespacios,
invariantes de un endomorfismo. 377
invariantes de una transformación ortogonal, 364
propios de un endomorfismo. 327
propios de una matriz, 330
propios generalizados, 402
suplementarios, 145
Subestructura. 735
Subgrupos, 747
Submatriz. 60
Sucesión indefinida, 731
Suma.
de ángulos, 772
de aplicaciones lineales, 176
de matrices, 45. 177
de subespacios, 139, 140

Suma *(Cont.)*
de subespacios unidimensionales, 144
de variedades lineales, 520
de vectores, 23, 116
directa, 141, 143
Suplementario ortogonal, 278
Supremo de un conjunto, 726
Sustituciones, 76, 752

Tangentes,
a las circunferencias, 449
a las cónicas, 562, 565, 606, 612
a las cuádricas, 587, 640
a una esfera, 482
Teorema,
contrario, 701
contrarrecíproco, 701
de Berstein, 733, 780
de Cantor, 733, 782, 786
de Cayley-Hamilton, 416
de Ceva, 540.
de Desargues, 540
de Gödel, 780
de isomorfismo, 187
de la base incompleta, 131, 274
de la dimensión, 129
de Menelao, 540
de Pitágoras, 268, 428
de Rouche, 100, 102, 189
de Zermelo, 784
directo, 701
espectral, 357
recíproco, 701
Teoremas, 691, 700, 783
Teoría de conjuntos, 781
Tesis, 700

Transformación ortogonal. 287
directa, 293
inversa, 293
Transformaciones, 712
ortogonales bidimensionales, 297
ortogonales tridimensionales, 300
Trasposiciones, 76, 753
Triangularización ortogonal de un endomorfismo, 381

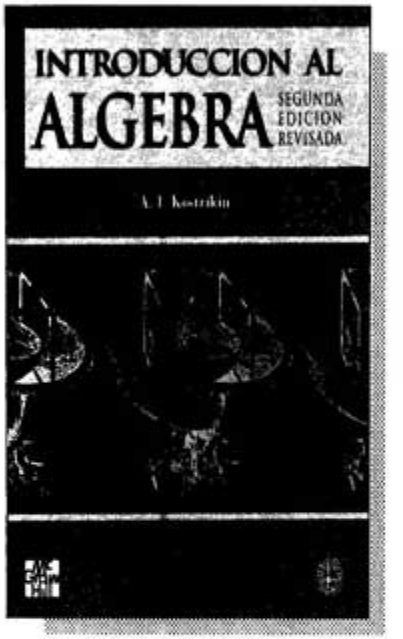
Ultimo teorema de Fermat, 778
Unión de conjuntos, 706

Valores propios, 327, 330
Variedad,
afín, 186
engendrada por una familia de puntos, 513
lineal, 509, 514, 529
afín, 509
de dimensión finita, 514
perpendicular a otra, 529
Variedades lineales ortogonales, 512
Vector, 116
columna, 24
de *n* componentes, 23
fila, 24
libre, 499
nulo, 25, 116
opuesto de otro, 25, 116
unitario, 267
Vectores,
conjugados, 225
libres, 721
ortogonales, 267
ortonormales, 267
propios, 327, 330
Volumen de un tetraedro, 480

INTRODUCCION AL ALGEBRA

2ª edición revisada

Alexei I. **KOSTRIKIN**
Universidad de Lomonósov. Moscú



440 pp.
ISBN: 84-7615-914-5

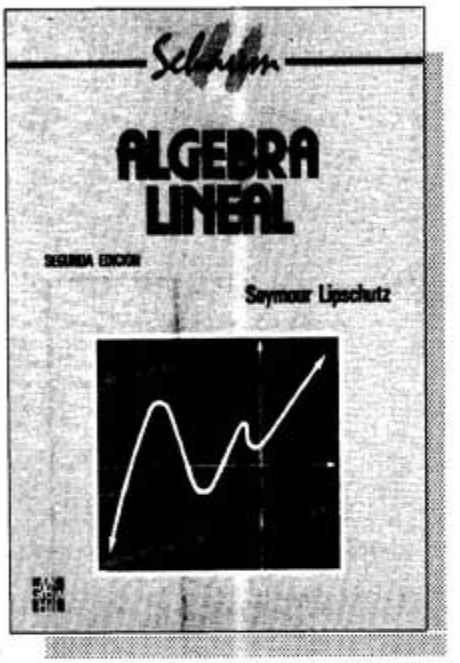
Uno de los objetivos de este libro es suministrar los conceptos y resultados necesarios a nivel introductorio, si bien el tratamiento es altamente formal y con un gran rigor matemático.

Los ejercicios juegan un papel importante. Tienen un carácter perfectamente teórico y sirven para desarrollar el tema principal. En muchos casos hay referencias en el texto sobre ellos, pero, pese a todo, están provistos de indicaciones detalladas para su resolución.

ALGEBRA LINEAL 2ª edición

Serie Schaum

Seymour **LIPSCHUTZ**



564 pp.
ISBN: 84-7615-758-4

Este libro de la serie *Schaum*, tiene por objetivo presentar una introducción al álgebra lineal que sea útil para todos los lectores, cualquiera que sea su campo de especialización, incluyendo más material del que generalmente se ve en un primer curso, con objeto de hacerlo más flexible.

Cada capítulo comienza con enunciados claros de las correspondientes definiciones, principios y teoremas. A estos siguen colecciones graduadas de problemas resueltos y problemas suplementarios.

UNA OBRA IMPRESCINDIBLE PARA PRACTICAR EL ALGEBRA LINEAL

OTRAS OBRAS DE INTERES PUBLICADAS POR McGRAW-HILL/INTERAMERICANA

ABELLANAS/GALINDO. *Métodos de Cálculo* (Schaum).
AYRES. *Algebra moderna* (Schaum).
AYRES. *Cálculo diferencial e integral* (3.ª ed.) (Schaum).
CHURCHILL. *Variable compleja y aplicaciones* (5.ª ed.).
GRANERO. *Algebra y geometría analítica.*
GRANERO. *Cálculo.*
GROSSMAN. *Algebra lineal con aplicaciones* (4.ª ed.).
KOSTRIKIN. *Introducción al álgebra* (2.ª ed.).
LARSON. *Cálculo y geometría analítica* (3.ª ed.).
LIPSCHUTZ. *Algebra lineal* (2.ª ed.) (Schaum).
PITA. *Algebra lineal.*
REES/SPARKS/SPARKS REES. *Algebra* (10.ª ed.).
RUDIN. *Análisis real y complejo.*
SPIEGEL/ABELLANAS. *Fórmulas y tablas de matemática aplicada* (Schaum).

9 788448 101343

ISBN: 84-481-0134-0